



**FACULTAD DE INGENIERÍA ARQUITECTURA Y
URBANISMO**

ESCUELA PROFESIONAL DE INGENIERÍA CIVIL

TESIS

**DESARROLLO DE UN MODELO DE REDES
NEURONALES ARTIFICIALES PARA PREDECIR
LAS PROPIEDADES MECÁNICAS DEL SUELO**

**PARA OPTAR EL TÍTULO PROFESIONAL DE:
INGENIERO CIVIL**

Autor:

Bach. Valderrama Purizaca, Frank Jesús

<https://orcid.org/0000-0002-2701-1151>

Asesor:

Dr. Muñoz Pérez, Sócrates Pedro

<https://orcid.org/0000-0003-3182-8735>

Línea de Investigación:

Infraestructura, Tecnología y Medio Ambiente

Pimentel – Perú

2020

**“DESARROLLO DE UN MODELO DE REDES NEURONALES ARTIFICIALES
PARA PREDECIR LAS PROPIEDADES MECÁNICAS DEL SUELO”**

Aprobación del jurado

Dr. Muñoz Pérez, Sócrates Pedro

ASESOR

Mg. Marín Bardales, Noé Humberto

PRESIDENTE

Mg. Villegas Granados Luis Mariano

SECRETARIO

Mg. Idrogo Pérez César Antonio

VOCAL

Dedicatoria

A mis padres, Marisol y Francisco, por su amor y apoyo infinito, cuyas palabras de aliento resuenan en mis oídos. A mis hermanos, Erick y Ronald, que son mi motivación.

Agradecimientos

En primer lugar, me gustaría agradecer a mi asesor de tesis, Dr. Sócrates Pedro Muñoz Pérez, quien fuese también mi profesor y director de escuela, cuya experiencia fue inestimable a la hora de elaborar esta investigación. Sus perspicaces comentarios me ayudaron a afinar mis ideas y a llevar mi trabajo a un nivel superior. Asimismo, por incentivarnos a realizar investigación y formar parte de este hermoso mundo.

Quisiera agradecer a mi amigo, Daniel Chávez, con quien, desde segundo año de carrera, comenzamos a realizar exposiciones nunca antes vistas. También quiero agradecer a los docentes de la Escuela Profesional de Ingeniería Civil de la Universidad Señor de Sipán por su valiosa formación a lo largo de mis estudios.

Quiero agradecer a mis padres por sus consejos brindados y el apoyo que me dieron durante los momentos difíciles. A mis hermanos, los mejores del mundo. Siempre estaré agradecido con todo lo que han hecho. Decirles que parece un sueño estar en esta etapa de la carrera, miro atrás y me doy cuenta de que la espera ha valido la pena.

A mis primos, Alexander, Cristhian, Renato, Nicol, Noemí, Joel... A toda mi familia, en especial a Emanuel y Mirella, que aún no saben lo que es una tesis, pero que han sido importantes. A Evelyn, Francisco y Milagros. También agradecer a una gran amiga, Mireya Troyes.

Un agradecimiento a todos aquellos con los que compartí esta casa de estudio, Bautista, Jennifer, León, Rosario, Elmer, Renzo, Gallardo, Lucy, Paola, Mario; y a unos grandes amigos, Andrea y Crosby, con quienes empezó todo. Fue divertido pasar tiempo con ustedes. GRACIAS.

Resumen

En este estudio se desarrollan modelos de redes neuronales artificiales para la predecir las propiedades mecánicas del suelo de la región Lambayeque. La tesis se realizó para facilitar la obtención de resultados precisos y se calculen en el menor tiempo posible. Se reunieron un total de 285 resultados de pruebas de proyectos de carreteras realizados en la región Lambayeque. Asimismo, las variables de entrada utilizadas fueron de los ensayos de contenido de humedad, contenido de sales, análisis granulométrico y límites de consistencia. Los modelos alcanzaron un error cuadrático medio favorable cuando estimaron la máxima densidad seca (0.003 g/cm^3), el óptimo contenido de humedad (2.42%), la relación de soporte de California al 95 % de la máxima densidad seca (6.35 %) y la relación de soporte de California al 100 % de la máxima densidad seca (21.80 %). En cuanto al error absoluto medio, se obtuvieron resultados favorables al predecir la máxima densidad seca (0.035 g/cm^3), el óptimo contenido de humedad (1.13 %), la relación de soporte de California al 95 % de la máxima densidad seca (1.65 %) y la relación de soporte de California al 100 % de la máxima densidad seca (2.79 %). Aunque son numerosos los elementos que influyen en el comportamiento del suelo, los resultados recogidos revelaron el rendimiento y las capacidades de las redes neuronales artificiales para predecir los atributos del suelo. Los modelos propuestos pueden utilizarse en proyectos reales como ayuda si se quiere obtener estas propiedades de manera rápida.

Palabras clave: parámetros de compactación, perceptrón multicapa, propiedades geotécnicas, propiedades índice del suelo, relación de soporte de California

Abstract

In this study, artificial neural network models are developed to predict the mechanical properties of the soil in the Lambayeque region. The thesis was done to facilitate the obtaining of accurate and calculated results in the shortest possible time. A total of 285 test results were collected for road projects carried out in the Lambayeque region. Moisture content, salt content, granulometric analysis, and consistency restrictions were also employed as input factors. The models achieved a favorable mean square error when estimating the maximum dry density (0.003 g/cm³), the optimum moisture content (2.42 %), the California stand ratio at 95 % of the maximum dry density (6.35 %), and the California stand ratio at 100 % of the maximum dry density (21.80 %). As for the mean absolute error, favorable results were obtained when predicting the maximum dry density (0.035 g/cm³), the optimum moisture content (1.13 %), the California support ratio at 95 % of the maximum dry density (1.65 %), and the California support ratio at 100 % of the maximum dry density (2.79 %). Even though numerous elements influence soil behavior, the findings collected revealed the performance and capabilities of artificial neural networks to predict soil attributes. The proposed models can be used in real projects as an aid if these properties are to be obtained quickly.

Keywords: compaction parameters, multilayer perceptron, geotechnical properties, index soil properties, California bearing ratio

ÍNDICE

Aprobación del jurado	ii
Dedicatoria	iii
Agradecimientos	iv
Resumen	v
Abstract	vi
I. INTRODUCCIÓN	13
1.1. Realidad problemática	13
1.2. Antecedentes de estudio.....	17
1.3. Teorías relacionadas al tema	20
1.4. Formulación del problema.....	53
1.5. Justificación e importancia del estudio	53
1.6. Hipótesis	54
1.7. Objetivos.....	54
1.7.1. Objetivo general.....	54
1.7.2. Objetivos específicos	54
II. MATERIAL Y MÉTODO	55
2.1. Tipo y diseño de investigación	55
2.2. Población y muestra.....	56
2.3. Variables y operacionalización	58
2.4. Técnicas e instrumentos de recolección de datos, validez y confiabilidad.....	62
2.5. Procedimiento de análisis de datos	67
2.6. Criterios éticos	69
2.7. Criterios de rigor científico.....	70

III. RESULTADOS.....	71
3.1. Resultados en tablas y figuras	71
3.2. Discusión de resultados	121
IV. CONCLUSIONES Y RECOMENDACIONES	124
4.1. Conclusiones	124
4.2. Recomendaciones	125
Referencias.....	126
ANEXOS.....	136

ÍNDICE DE FIGURAS

Figura 1	Tamices utilizados en un análisis granulométrico.....	21
Figura 2	División del suelo según los límites de Atterberg	22
Figura 3	Desarrollo y expansión de la inteligencia artificial	26
Figura 4	Curvas de error idealizadas para modelos de ANN generales	28
Figura 5	Estructura de una neurona biológica.....	31
Figura 6	Estructura de una red neuronal artificial.....	32
Figura 7	Perceptrón sin función de activación.....	34
Figura 8	Perceptrón con función de activación.....	35
Figura 9	Arquitectura general de un perceptrón multicapa	36
Figura 10	Esquema básico de una neurona artificial.....	37
Figura 11	Función de activación Lineal.....	40
Figura 12	Función de activación Sigmoide.....	41
Figura 13	Función de activación ReLU	42
Figura 14	Función de activación Tangente hiperbólica	43
Figura 15	Visualización de la validación cruzada	47
Figura 16	Procedimiento seguido para el desarrollo de las ANN.....	64
Figura 17	Distribución del conjunto de datos según su clasificación USCS.....	72
Figura 18	Distribución del conjunto de datos según su clasificación AASTHO	73
Figura 19	Arquitectura inicial de los modelos de ANN con 8 variable de entrada.....	76
Figura 20	Arquitectura inicial de los modelos de ANN con 10 variables de entrada.....	76
Figura 21	Diagrama de dispersión para MDD y las variables de entrada	78
Figura 22	Diagrama de dispersión para OMC y las variables de entrada	79
Figura 23	Diagrama de dispersión para el CBR al 95 % de la MDD y las variables de entrada.....	80

Figura 24	Diagrama de dispersión para el CBR al 100 % de MDD y las variables de entrada.....	81
Figura 25	Diagrama de dispersión para el CBR y la compactación	82
Figura 26	Predicciones realizadas por la ANN MARK 1	90
Figura 27	Predicciones realizadas por la ANN MARK 3	92
Figura 28	Rendimiento de los modelos de ANN desarrollados para mejorar el MARK 1	94
Figura 29	Rendimientos de los modelos desarrollados para mejorar el MARK 1 según el número de neuronas ocultas.....	95
Figura 30	Predicciones realizadas por la ANN MARK 2	100
Figura 31	Predicciones realizadas por la ANN MARK 4	102
Figura 32	Rendimiento de los modelos de ANN desarrollados para mejorar el MARK 2.....	104
Figura 33	Rendimientos de los modelos desarrollados para mejorar el MARK 2 según el número de neuronas ocultas.....	105
Figura 34	Predicciones realizadas por la ANN MARK 5	109
Figura 35	Rendimiento de los modelos de ANN desarrollados para mejorar el MARK 5.....	111
Figura 36	Rendimientos de los modelos desarrollados para mejorar el MARK 5 según el número de neuronas ocultas.....	112
Figura 37	Predicciones realizadas por la ANN MARK 6	116
Figura 38	Rendimiento de los modelos de ANN desarrollados para mejorar el MARK 6.....	118
Figura 39	Rendimientos de los modelos desarrollados para mejorar el MARK 6 según el número de neuronas ocultas.....	119

ÍNDICE DE TABLAS

Tabla 1	Símbolos utilizados en el sistema USCS.....	24
Tabla 2	Tipos de hiperparámetros en modelos de aprendizaje automático	30
Tabla 3	Resumen de la información recopilada del SEACE	57
Tabla 4	Operacionalización de la variable independiente	60
Tabla 5	Operacionalización de la variable dependiente	61
Tabla 6	VARIABLES consideradas para los modelos de ANN.....	65
Tabla 7	Información recopilada del SEACE de los años 2019, 2020 y 2021	74
Tabla 8	Entradas y Salidas utilizadas	75
Tabla 9	Entradas y salidas de modelos según literatura	77
Tabla 10	Estadísticas del conjunto de entrenamiento y prueba para las variables de entrada.....	83
Tabla 11	Estadísticas del conjunto de entrenamiento y prueba para las variables de salida	84
Tabla 12	Prueba U de Mann-Whitney para la igualdad de medias entre los conjuntos de entrenamiento y prueba para las variables de entrada	85
Tabla 13	Prueba U de Mann-Whitney para la igualdad de medias entre los conjuntos de entrenamiento y prueba para las variables de salida.....	86
Tabla 14	Mejores 10 ANN para predecir el CBR al 100% de la MDD (8 entradas).....	87
Tabla 15	Mejores 10 modelos para predecir el CBR al 100% de la MDD (10 entradas).....	88
Tabla 16	Hiperparámetros del modelo de ANN MARK 1	89
Tabla 17	Hiperparámetros del modelo de ANN MARK 3.....	91
Tabla 18	Arquitectura e hiperparámetros de los modelos de ANN desarrollados para mejorar el MARK 1	93

Tabla 19	Estructura y rendimiento de los modelos de ANN desarrollados para predecir el CBR al 100 % de MDD	96
Tabla 20	Mejores 10 ANN para predecir el CBR al 95% de la MDD (8 entradas).....	97
Tabla 21	Mejores 10 modelos para predecir el CBR al 95% de la MDD (10 entradas).....	98
Tabla 22	Hiperparámetros del modelo de ANN MARK 2.....	99
Tabla 23	Hiperparámetros del modelo de ANN MARK 4.....	101
Tabla 24	Arquitectura e hiperparámetros de los modelos de ANN desarrollados para mejorar el MARK 2.....	103
Tabla 25	Estructura y rendimiento de los modelos de ANN desarrollados para predecir el CBR al 95 % de MDD	106
Tabla 26	Mejores 10 modelos para predecir la MDD con 8 variables de entrada.....	107
Tabla 27	Arquitectura e hiperparámetros de los mejores modelos aplicando optimización bayesiana	108
Tabla 28	Arquitectura e hiperparámetros de los modelos de ANN desarrollados para mejorar el MARK 5.....	110
Tabla 29	Estructura y rendimiento de los modelos de ANN desarrollados para predecir MDD.....	113
Tabla 30	Mejores 10 modelos para predecir el OMC con 8 variables de entrada.....	114
Tabla 31	Arquitectura e hiperparámetros de los mejores modelos aplicando optimización bayesiana	115
Tabla 32	Arquitectura e hiperparámetros de los modelos de ANN desarrollados para mejorar el MARK 6.....	117
Tabla 33	Estructura y rendimiento de los modelos de ANN desarrollados	120

I. INTRODUCCIÓN

1.1. Realidad problemática

Los suelos están compuestos de grava, arena, limo y arcilla como resultado de la desintegración, depósito y transporte de rocas. Es necesario conocer su comportamiento ante determinadas cargas para garantizar la vida útil de las estructuras. Por esta razón, es importante estimar las propiedades del suelo de manera precisa. Para ello existen diversos métodos que permiten examinar los suelos, los mismos que se complementan entre sí (Özyankı, 2019).

La resistencia característica del suelo es importante para el buen desempeño de un pavimento. Se determina los requerimientos mínimos y se evalúa si el material existente es el adecuado (Al-Busultan *et al.*, 2020; Sandoval Vallejo & Rivera Mena, 2019; Verma & Kumar, 2020). El ensayo de la Relación de Soporte de California (CBR) utiliza un índice para ayudar a determinar la resistencia del suelo (Ministerio de Transportes y Comunicaciones [MTC], 2016). Con un CBR bajo, el espesor del pavimento será mayor al de un suelo con un CBR alto (Katte *et al.*, 2019).

1.1.1. Nivel internacional

Ahora bien, para que un país pueda prosperar y desarrollarse es necesario que cuente con adecuadas infraestructuras. El número de redes de carreteras, por ejemplo, son un indicador de progreso económico, social y comercial de un país. Teniendo en cuenta esto, resulta evidente de porqué Camerún asigna el 28 % de los fondos del Banco Mundial al sector transporte (Yato Katte *et al.*, 2018).

En busca de este desarrollo, desde hace años, los ingenieros han diseñado pavimentos empleando diversos métodos numéricos y siguiendo los criterios establecidos en los reglamentos. Aunque, para ello, se necesitaba conocer una serie de parámetros del suelo que son complicados de obtener (Özgan *et al.*, 2015). Esto debido a que los suelos tienen diversas propiedades que influyen en su comportamiento (Özyankı, 2019).

Se ha investigado la influencia de las características físicas del suelo en la compactación y el CBR en busca de una técnica sencilla para determinar estos valores. (Al-Busultan *et al.*, 2020; Khuntia *et al.*, 2015). La distribución de las partículas, la gravedad específica, el límite líquido (LL), el límite plástico (PL), el límite de contracción (SL) y el índice de plasticidad (PI), son algunas de las características más estudiadas para este fin (Tizpa *et al.*, 2015).

Desde entonces, muchos investigadores han propuesto diversas ecuaciones que permiten, de una u otra manera, predecir los parámetros de compactación y el CBR de los suelos según sus propiedades físicas (Khalid & Rehman, 2018). Aunque en la actualidad se sigue investigando con el fin de encontrar modelos que permitan estimar los valores de estas propiedades del suelo (Ardakani & Kordnaeij, 2019).

Cabe destacar que casi todas las correlaciones usadas se realizan haciendo un análisis de regresión, los cuales requieren parámetros de entrada que son propensos a incertidumbres e inexactitudes. Como muchos factores son los que afectan a las propiedades del suelo, los resultados pueden presentar variaciones que, de ser usados para el diseño de pavimentos, podrían dar valores poco semejantes a la realidad (Ardakani & Kordnaeij, 2019).

El desarrollo de Redes Neuronales Artificiales (ANN) es una alternativa para predecir las distintas propiedades de los suelos. Las ANN reducen el tiempo y el coste, además, presentan un alto rendimiento a comparación de los modelos empíricos (Taleb Bahmed *et al.*, 2019). Siendo esas las razones para que estos modelos se apliquen en diversos campos y se consideren métodos muy eficientes (Ardakani & Kordnaeij, 2019).

A pesar de que el ensayo de CBR tiene una gran aceptación y es muy utilizado para el diseño de pavimentos, no siempre es posible realizarlo (Sandoval Vallejo & Rivera Mena, 2019). Entre las principales razones está que determinar el CBR no es una tarea sencilla, es costoso y el tiempo consumido para realizarlo es grande (Rehman *et al.*, 2017). A esto se le añade que el tiempo disponible para realizar los ensayos en etapas previas al proyecto es limitado, además, que existen limitaciones económicas para caracterizar los materiales (Bogado *et al.*, 2017).

1.1.2. Nivel nacional

En cuanto al desarrollo de Perú, el sector de la construcción experimentó un descenso de la producción del 67,2 % en el segundo trimestre de 2020. Esto se debió a que los proyectos de construcción, tanto gubernamentales como privados, se detuvieron debido a la pandemia. La inversión en infraestructuras de transporte también disminuyó afectando a la actividad productiva del Perú, que ya para mitad del año 2020, había disminuido en un 17.3 % a comparación del año 2019 (Instituto Nacional de Estadística e Informática [INEI], 2020).

Una manera de contribuir con el progreso del país es invirtiendo en proyectos de construcción de buena calidad. Esto no sucede en la zona tropical del Perú, pues las carreteras y aeropuertos que se construyen presentan graves problemas geotécnicos. La causa principal es la complejidad de las variables que afectan al comportamiento del suelo. Debido a factores como la composición química y la estructura morfológica del suelo, los suelos de la selva central del norte de Perú pueden dividirse en hasta cinco grupos distintos de rocas (Carrillo gil, 2017).

Por esta razón, los suelos de la región tropical del Perú poseen una humedad y límites de consistencia muy variados. A una profundidad de 3 metros desde la superficie, se pueden encontrar desde suelos completamente secos hasta suelos con un 100 % de humedad natural. El porcentaje de humedad natural promedio del suelo a una profundidad de 40 metros varía entre el 20 % y el 30 %. A unos 50 metros de profundidad se pueden encontrar suelos con un LL que varía entre 20 % y 70 %, llegando en algunos casos hasta el 90 %; el PL varían entre 16 % y 32 % a 40 % (Carrillo gil, 2017).

Por otro lado, en la selva alta del Perú se pueden encontrar suelos residuales provenientes de rocas duras cuya máxima densidad seca (MDD) está entre 1.85 g/cm³ y 2.25 g/cm³ y óptimo contenido de humedad (OMC) está de 5 % a 14 %. En la selva baja, cuyos suelos son originados por rocas blandas, los suelos poseen una MDD entre 1.50 g/cm³ a 1.85 g/cm³ y un OMC entre 9 % a 21 % (Carrillo gil, 2017).

En cuanto a la facilidad de conocer las propiedades de los suelos, un ensayo de CBR puede costar alrededor de S/530 (Universidad Nacional de Ingeniería, 2019). Así pues, realizar un ensayo de CBR en cada una de las muestras adquiridas

de las calicatas resultaría muy costoso si se quiere conocer el comportamiento exacto del suelo a lo largo de una sección de estudio.

1.1.3. Nivel local

La disminución de obras de carreteras es un factor que ha afectado en el desarrollo de la región Lambayeque cuya producción, para el segundo trimestre de 2020, bajó en un 24.4 % influenciado en un 65.5 % por el sector Construcción. Del mismo modo, el sector Construcción disminuyó en un 64.5 %, ya que la pandemia frenó el progreso físico de los proyectos de los gobiernos locales (INEI, 2020).

Para ayudar con el desarrollo de la región, es necesario construir carreteras que cumplan con los requerimientos mínimos. Por ejemplo, realizar 6 calicatas por kilómetro para carreteras con un alto volumen de tránsito y 1 por kilómetros para carreteras con un volumen de tráfico bajo (MTC, 2014). Asimismo, se debe tener entre 16 kg y 29 kg de suelo para un ensayo de Proctor Modificado, debiéndose extraer en campo entre 23 kg y 45 kg de suelo, aproximadamente (Manual de Ensayo de Materiales, 2017, pág. 108).

El Manual de carreteras (2014) establece un número mínimos de ensayos de CBR a realizar. Por ejemplo, 1 ensayo cada kilómetro para carreteras con un alto volumen de tránsito y 1 ensayo de CBR cada 3 kilómetros para carreteras con un bajo volumen de tránsito (pág. 30). Como no se puede tener una representación detallada del comportamiento del suelo, la carretera se divide en sectores con la finalidad de encontrar tramos homogéneos según las características del suelo y así obtener un CBR representativo, el cual se usa para diseñar el pavimento.

En resumen, los ensayos de laboratorio son muy beneficiosos para los ingenieros geotécnicos (Taleb Bahmed *et al.*, 2019). Sin embargo, las propiedades del suelo varían mucho y si se toma en cuenta el número de calicatas con el número de ensayos, resulta costoso sacar una cantidad de muestras que permita determinar las características del suelo en su totalidad (MTC, 2014). Asimismo, el tiempo consumido para elaborar los ensayos es elevado (4 días para el CBR, por ejemplo) cuando se trata de proyectos grandes (MTC, 2016).

A diferencia de otros países, donde las ANN están siendo aplicadas en diversas ramas de la ingeniería civil, en la región Lambayeque son muy escasas

las investigaciones (Valderrama Purizaca *et al.*, 2020). En general, se usan para predecir propiedades como MDD, OMC y CBR, entre otras (Taleb Bahmed *et al.*, 2019). La rapidez con la que estos modelos producen resultados, su capacidad de aprender y mejorar son unas de sus ventajas (Khalid & Rehman, 2018). Siendo conveniente desarrollar modelos de ANN que permitan predecir propiedades del suelo (Tizpa *et al.*, 2015).

1.2. Antecedentes de estudio

Este apartado tiene la finalidad de presentar el avance que ha tenido el tema de investigación durante los últimos 7 años. Por ello, se hizo una búsqueda de artículos científicos e investigaciones realizadas y, posteriormente, se hizo el respectivo análisis. Se describen los objetivos, metodologías, resultados y conclusiones a las que llegaron. Investigaciones relacionadas con la predicción de parámetros como MDD, OMC y CBR aplicando modelos de ANN.

1.2.1. Nivel internacional

Tenpe y Patel (2020), en su artículo “Aplicación de la programación de expresión genética y redes neuronales artificiales para la predicción del CBR”, desarrollaron modelos de ANN para predecir el CBR del suelo. Se encontró que las mejores predicciones se hicieron con el porcentaje de grava, porcentaje de arena, PI, OMC y MDD como entradas. Aunque, también se hicieron pruebas con el porcentaje de finos, LL y PL. Según los resultados, el porcentaje de grava y arena son los parámetros más significativos con un 26.83 % y 23.37%, respectivamente.

Alam *et al.* (2020), quienes en su artículo “Predicción del CBR de suelos finos de la cuenca de Bengala aplicando programación de expresión genética, redes neuronales artificiales y el método de Krigging”, desarrollaron modelos para predecir el CBR de los suelos. Las variables de entrada usadas en los modelos fueron el coeficiente de uniformidad, la gravedad específica, el coeficiente de curvatura, LL, PL, PI, OMC y MDD. Al finalizar su investigación, los resultados mostraron que los modelos propuestos predicen el CBR del suelo con mucha precisión.

Rajakumar y Reddy babu (2020), en su artículo “Estudio experimental y modelado de redes neuronales de subrasantes expansivas estabilizadas con subproductos de residuos industriales y geomalla”, desarrollaron modelos de regresión lineal y ANN para predecir el CBR de los suelos. Utilizaron 6 variables en la capa de entrada y emplearon el algoritmo de Levenberg-Marquardt. La ANN estimó el CBR con un coeficiente de correlación (R) de 0.94317 concluyendo que el modelo de ANN es más preciso que el de regresión lineal.

Fikret Kurnaz y Kaya (2019), en su investigación “Uso de una red neuronal de tipo GMDH para predecir el valor de soporte de California (CBR) de los suelos compactados”, tuvieron como objetivo predecir el CBR utilizando el Método de Manejo de Datos en Grupo. Las variables de entrada fueron el porcentaje de grava, el porcentaje de arena, el porcentaje de finos, el OMC y la MDD. Los resultados obtenidos demuestran que el modelo GMDH predice mejor los valores de CBR a comparación del análisis de regresión múltiple.

Ghorbani y Hasanzadehshooilli (2018), en su artículo “Uso de modelos ANN y EPR para predecir el UCS y el CBR de arenas limosas sulfatadas estabilizadas con microsilíce; aplicación a la mezcla de suelos profundos”, se emplearon 90 pruebas de resistencia a la compresión confinada y CBR de arenas del desierto de Irán para desarrollar un modelo de ANN y Regresión Evolutiva de Polinomios. El coeficiente de determinación (R^2) de la ANN con arquitectura 5-5-8-1 fue de 0.9917 y para las arquitecturas de 4-2-3-1, 4-2-3-1 y 4-5-1 fue de 0.99; las cuales predijeron con bastante éxito.

Suthar y Aggarwal (2018), en su estudio “Predicción del valor de CBR de cenizas de estanque estabilizadas con cal y residuos de cal usando modelos ANN y RM”, tuvieron como objetivo predecir el CBR aplicando ANN y regresión múltiple. Para ello, utilizaron 51 ejemplos de CBR obtenidos de ensayos de laboratorio. Las métricas usadas para evaluar los modelos fueron la raíz del error cuadrático medio (RMSE), el coeficiente de correlación (R) y el error absoluto medio (MAE). Se pudo observar que ambos modelos predicen el CBR con mucha exactitud (R igual a 0.96), además, que el periodo de curado afecta significativamente al CBR de las cenizas.

Erzin y Turkoz (2016), en su investigación “Utilización de redes neuronales para predecir el valor de CBR de determinadas arenas del Egeo”, realizaron

ensayos de CBR a las arenas de 9 lugares distintos de la ciudad de Manisa, Turquía, para desarrollar un modelo de regresión múltiple y ANN. Las variables de entrada fueron: gravedad específica, coeficiente de uniformidad, coeficiente de curvatura, densidad seca, contenido de humedad, porcentaje de cuarzo, feldespatos, calcita, corindón, y otros minerales. La ANN con arquitectura 10-5-1 (R^2 de 0.9783) y la regresión múltiple (R^2 de 0.8120) son exactas.

Albusoda *et al.* (2020), en su investigación “Densidad seca basada en las propiedades índices del suelo mediante el uso de un sistema experto”, proponen un modelo de ANN para predecir la MDD del suelo. Se recopilaron 99 muestras de suelo y se utilizaron sus resultados de ensayos de laboratorio. Los modelos desarrollados tuvieron diferentes números de variables de entrada (9, 6, 5 y 3) y una capa oculta de 10 neuronas. La ANN dio resultados muy similares a los esperados con un R^2 igual a 0.97368; llegando a la conclusión que se puede aplicar el modelo de ANN para predecir esta propiedad.

Ardakani y Kordnaeij (2019), en su investigación “Predicción de las características de compactación del suelo mediante una red neuronal de tipo GMDH y un algoritmo genético”, tuvieron como objetivo estimar los parámetros de compactación según las propiedades índices del suelo. Para ello, utilizaron un total de 212 ejemplos. Las variables de entrada fueron el contenido de finos, contenido de arena, LL y PL. Los resultados de los valores medidos y los valores predichos fueron muy similares por lo que se llegó a la conclusión que los parámetros más influyentes en la compactación son el límite líquido y PL.

1.2.2. A nivel nacional

Manrique Dávila (2019), en su tesis “Predicción de propiedades mecánicas de suelos utilizando redes neuronales artificiales”, predijo el CBR y los parámetros de compactación del suelo. Se usó 212 muestras obtenidas de proyectos de carreteras de la base de datos de Provías Nacional. Se utilizaron como entradas el porcentaje de grava, el porcentaje de arena, el porcentaje de finos, LL, PL, PI. Se predijo el CBR al 100 % (R^2 de 0.7063), el CBR al 95 % (R^2 de 0.8232), el OMC (R^2 de 0.8433) y la MDD (R^2 de 0.9245).

1.2.3. A nivel local

En la región Lambayeque no se han realizado investigaciones en las cuales utilicen modelos de ANN para predecir las propiedades del suelo. Sin embargo, se han encontrado investigaciones donde las ANN son utilizadas para predecir la resistencia del concreto y otras investigaciones donde utilizan modelos de regresión para correlacionar las propiedades del suelo y encontrar ecuaciones que permitan predecir dichas propiedades.

1.3. Teorías relacionadas al tema

Este apartado proporciona los fundamentos teóricos de las propiedades de los suelos y, asimismo, sirve como base para la comprensión de los modelos de ANN. En esta parte se busca: (a) presentar una descripción de las principales propiedades geotécnicas del suelo; (b) comprender los conceptos fundamentales del Aprendizaje Profundo; y (c) conocer los principales tipos de ANN, funciones de activación, etapa de entrenamiento, etapa de validación y etapa de prueba, con algunas discusiones de sus ventajas y desventajas teóricas, a la hora de desarrollar modelos de ANN.

Para poder presentar estas teorías se realizó una revisión de artículos científicos y libros de alto impacto, cuyos autores son los referentes en cada tema presentado. Braja M. Das, Geoffrey Hinton, Goodfellow y François Chollet, por mencionar algunos. Por último, esta parte fue dividida de la siguiente manera: propiedades geotécnicas de los suelos, inteligencia artificial, redes neuronales artificiales y desarrollo de modelos de ANN.

1.3.1. Propiedades geotécnicas de los suelos

Comprender las propiedades de los suelos que rigen su comportamiento es necesaria para el diseño de los pavimentos. El ingeniero debe estar familiarizado con los fundamentos de la mecánica del suelo para determinar si las características son aceptables y tener en cuenta que el suelo no es homogéneo. Propiedades como el tamaño de los granos del suelo, límites de consistencia, compactación y

CBR pueden determinarse en el laboratorio. Sin embargo, existen ocasiones en la que esto no es posible (Braja M. Das & Nagaratnam Sivakugan, 2019).

1.3.1.1. Distribución del tamaño de partículas

Esta propiedad permite determinar la proporción de tamaño de partículas existentes en el suelo. Es necesaria para clasificar el suelo y comprender sus propiedades geotécnicas. Las cualidades geotécnicas de los suelos de grano grueso están influenciadas por el tamaño del grano del suelo. La plasticidad, en cambio, determinará el comportamiento de los suelos finos (Braja M. Das & Nagaratnam Sivakugan, 2019).

El análisis granulométrico por tamizado tiene como objetivo determinar cómo se distribuyen las partículas del suelo. Se usa una serie de tamices con diversas aberturas, como se ilustra en la Figura 1 (Braja M. Das & Nagaratnam Sivakugan, 2019). Los tamaños estándar que se utilizan son 3 in, 2 in, 1 ½ in, 1 in, ¾ in, ⅜ in, n.º 4, n.º 10, n.º 20, n.º 40, n.º 60, n.º 140, n.º 200. Como alternativa, cuando al dibujar la gradación se quiere tener un espaciado semejante entre puntos, se pueden utilizar los tamaños de 3 in, 1 ½ in, ¾ in, ⅜ in, n.º 4, n.º 8, n.º 16, n.º 30, n.º 50, n.º 100, n.º 200 (Manual de ensayo de materiales, 2017, pág. 44).

Figura 1

Tamices utilizados en un análisis granulométrico



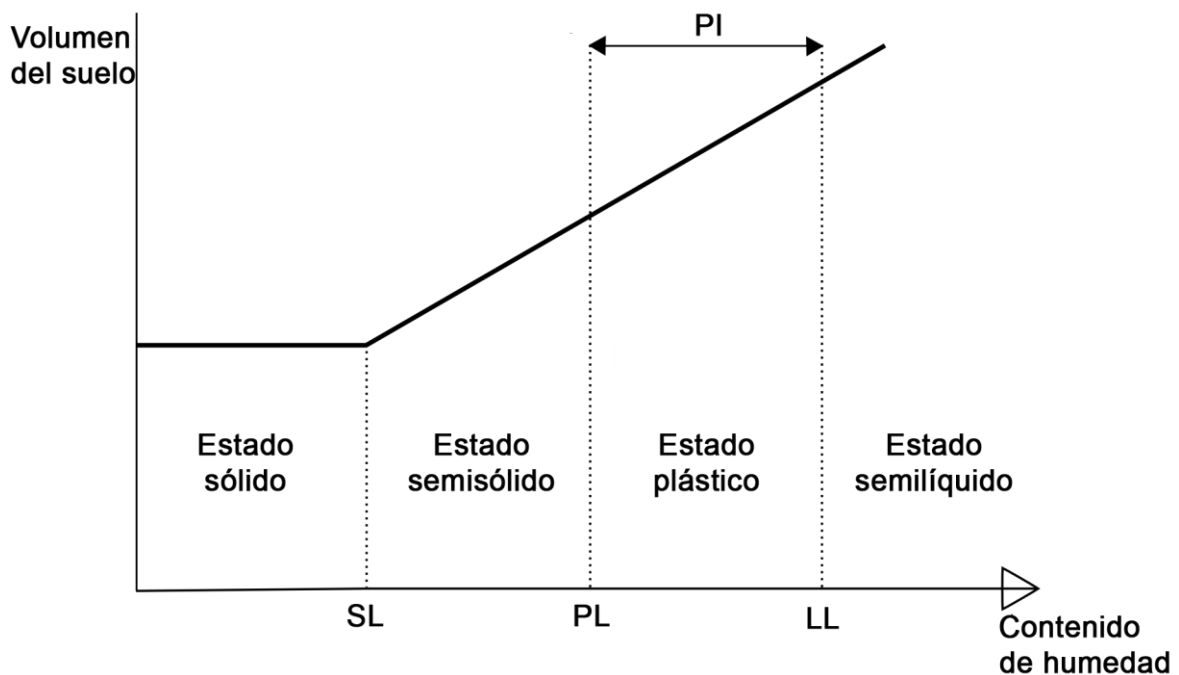
Nota. De *Test Sieves*, por rsaad77, 2018, Wikimedia Commons (https://commons.wikimedia.org/wiki/File:Test_Sieves.jpg). CC BY-SA 4.0.

1.3.1.2. Límites de Atterberg

Los límites de consistencia, o límites de Atterberg, son una de las propiedades que ayudan a determinar el comportamiento del suelo. Como se ilustra en la Figura 2, estos índices dividen el suelo como sólido, semisólido, plástico y semilíquido. Dado que el porcentaje de humedad de un suelo determina su comportamiento, un suelo arcilloso combinado con mucha agua fluirá y se comportará como un semilíquido; mientras que, un suelo seco tendrá un comportamiento plástico, semisólido o sólido (Braja M. Das & Nagaratnam Sivakugan, 2019).

Figura 2

División del suelo según los límites de Atterberg



Nota. Adaptado de *Atterbergse grenzen*, por MADe at Dutch Wikipedia, 2006, Wikimedia Commons (https://commons.wikimedia.org/wiki/File:Atterbergse_grenzen.PNG). CC BY-SA 2.0 BE.

Los límites de Atterberg fueron desarrollados por el científico sueco Atterberg, a principios de 1900 (Raad Al-Adhadh *et al.*, 2020). Son índices empíricos que determinan el contenido de agua del suelo. Estos a su vez, definen

los estados mecánicos del suelo y reflejan la capacidad de los suelos finos para resistir cargas externas (Zhou & Lu, 2021).

- **LL:** porcentaje de humedad donde en el que el suelo se transforma de estado semilíquido a plástico (Braja M. Das & Nagaratnam Sivakugan, 2019). El ensayo se realiza con la copa Casagrande y se determina cuando dos mitades de pastas de suelo, separadas $\frac{1}{2}$ ", se juntan luego de 25 golpes (MTC, 2016).
- **PL:** porcentaje de humedad donde en el que el suelo se transforma de estado plástico a semisólido (Braja M. Das & Nagaratnam Sivakugan, 2019). Se determina como la humedad con la que se pueden formar barritas de suelo sin que se desmoronen (MTC, 2016).
- **SL:** La cantidad de humedad necesaria para convertir un suelo semisólido en sólido (Braja M. Das & Nagaratnam Sivakugan, 2019).
- **PI:** La diferencia entre LL y PL se utiliza para calcular este valor (MTC , 2016).

1.3.1.3. Clasificación de los suelos

Existen dos sistemas usados para clasificar los suelos. Para empezar, el método AASTHO divide los suelos en ocho categorías: suelos granulares (grupos A-1, A-2 y A-3); suelos finos (grupos A-4, A-5, A-6 y A-7); y suelos orgánicos (grupo A-8). Comité de Clasificación de Materiales para Subrasantes y Carreteras de Tipo Granular del Consejo de Investigación de Carreteras desarrolló este método (Braja M. Das & Nagaratnam Sivakugan, 2019).

El USCS fue propuesto en 1942 por A. Casagrande. Como se ve en la Tabla 1, el método clasifica los suelos en grupos y subgrupos. Son suelos de grano grueso cuando más del 50% del suelo queda retenido en el tamiz n.º 200. Estos a su vez se clasifican en gravas (más del 50 % de suelo es retenido en el tamiz n.º 4) y arenas (el 50 % o más de suelo pasa el tamiz n.º 4). Los suelos de grano fino se identifican cuando el 50 % o más pasa el tamiz n.º 200 (Braja M. Das & Nagaratnam Sivakugan, 2019).

Tabla 1*Símbolos utilizados en el sistema USCS*

Grupo de suelo	Descripción
Suelo de grano Grueso:	
Grava (G)	Bien graduada (W)
Arena (S)	Mal graduada (P)
	Limoso (M)
	Arcillosos (C)
Suelo de grano Fino:	
Limo (M)	Baja plasticidad (L)
Arcilla (C)	Alta plasticidad (H)
Limos y arcillas orgánicas (O)	
Otros:	
Turba y suelo altamente orgánico (Pt)	—

Nota. Adaptado de *Principles of Foundation Engineering* (pág. 22) por Braja M. Das y Nagaratnam Sivakugan, 2019, Cengage Learning.

1.3.1.4. Compactación de los suelos

El ensayo para determinar estos valores son el Proctor estándar y el Proctor modificado. La compactación es el proceso mecánico de disminución del contenido de aire de los suelos. La compactación es una de las técnicas más antiguas y sencillas de realizar. Cuando se compacta cerca del OMC se consiguen mejores cualidades del suelo (Braja M. Das & Nagaratnam Sivakugan, 2019).

La compactación es una de las técnicas más eficientes y prácticas que mejora la capacidad de carga de un suelo (Özgan *et al.*, 2015). También permite calcular la proporción entre el peso específico seco y el contenido de agua (MTC, 2016). Asimismo, según Vafoeva (2020), es la forma más eficaz de aumentar las cualidades físicas y mecánicas del suelo, así como su capacidad de carga.

1.3.1.5. Relación de soporte de california (CBR)

Según Li *et al.* (2020), el CBR es uno de los parámetros geotécnicos que sirve para determinar el espesor adecuado de los pavimentos flexibles. De igual manera, Ji *et al.* (2020) señala que el CBR es un parámetro crucial para la construcción de estructuras de subrasante, pues al realizar un ensayo de CBR se conoce la capacidad de carga de la subrasante para resistir las deformaciones. Siendo así que, cuando la subrasante carece de capacidad portante significa que posee un bajo valor de CBR (Mina *et al.*, 2019).

Nini (2019) señala que la cantidad de agua que se filtra por debajo de una carretera como resultado de las lluvias tiene un impacto en el CBR del suelo. Por esta razón, es importante que el material que conforme la subrasante tenga la capacidad de drenar el agua. Si esto no ocurre, el pavimento puede verse afectado puesto que la resistencia del pavimento se alteraría si el porcentaje de humedad del material de la subrasante cambiara.

1.3.2. Inteligencia Artificial

La Inteligencia Artificial (AI) es una disciplina que su objetivo es emular en sistemas artificiales un conjunto de facultades intelectuales realizadas por las personas. Percepciones sensoriales como la visión, audición y reconocimiento de patrones (Benitez, 2014). En un principio, la AI era capaz de resolver problemas que eran complicados de resolver para una persona (Casas Roma *et al.*, 2019). El reto era resolver tareas intuitivas como reconocer la voz o identificar el rostro de alguien (Goodfellow *et al.*, 2016).

Para resolver estas dificultades se necesitaba que los modelos de AI tuvieran la capacidad de adquirir su propio conocimiento (Goodfellow *et al.*, 2016). Según Casas Roma *et al.* (2019), todo el conjunto de algoritmos y métodos que permiten que una máquina aprenda de manera automática o de experiencias pasadas, se le denomina Aprendizaje Automático (*Machine Learning*, en inglés). Estos algoritmos son capaces de aprender cuando su rendimiento (P) en cierta tarea (T) mejora con la experiencia (E) (Goodfellow *et al.*, 2016).

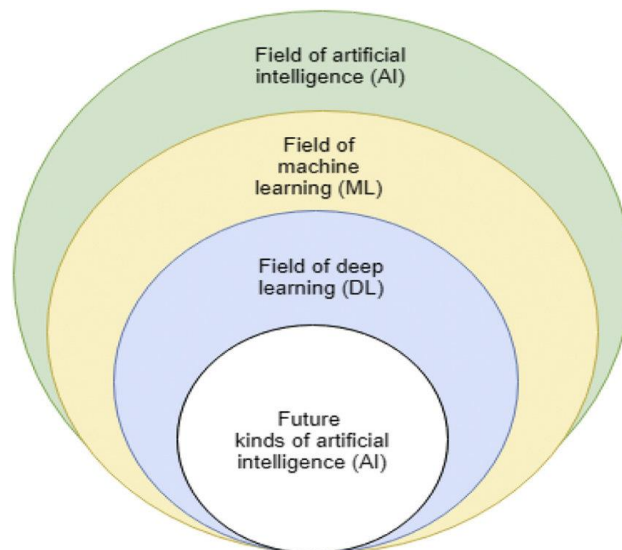
Ahora bien, con respecto a la manera de aprender, Alan Turing, matemático británico, quien anticipó futuros desarrollos de la AI e intuyó la importancia que

tendría el Aprendizaje Automático en el desarrollo de las mismas, mencionaba que era más factible que las máquinas primero emulen la mente de un niño y luego pasaran por un proceso de aprendizaje hasta llegar al equivalente de una mente adulta, en lugar de gastar esfuerzos haciendo que emularan la mente de un adulto (Meseguer Gonzalez & Badia, 2017).

En cuanto al Aprendizaje Profundo (*Deep Learning*, en inglés), es una rama del Aprendizaje Automático, como se ve en la Figura 3, que fue propuesto por Geoffrey Hinton en el año 2006 (Yingying Wang *et al.*, 2020). El Aprendizaje Profundo busca que el modelo sea capaz de representar conceptos complejos a partir de conceptos más simples. Por ejemplo, que una ANN pueda reconocer un coche, luego de aprender que las llantas, puertas, entre otros, son algunas partes del vehículo. Cuando se habla de aprendizaje profundo también se hace referencia a la profundidad del modelo, conocidas como capas (Casas Roma *et al.*, 2019).

Figura 3

Desarrollo y expansión de la inteligencia artificial



Nota. De “Un estudio sobre el estado del arte en las aplicaciones de las redes neuronales artificiales” por Abiodun *et al.*, 2018, *Heliyon*, 4(11), pág. 9. (<https://doi.org/10.1016/j.heliyon.2018.e00938>). CC BY 4.0.

Son tres los principales tipos de tareas que se resuelven: (a) la clasificación, cuando determinadas valores son asignadas a un conjunto, por ejemplo, el

reconocimiento de objetos (Casas Roma *et al.*, 2019); (b) la regresión, cuando se predice un valor numérico a partir de unas entradas determinadas (Goodfellow *et al.*, 2016); y (c) el agrupamiento, cuando un conjunto de valores es dividido en grupos según sus similitudes para crear un modelo que asigne nuevas entradas a estos grupos ya formados (Casas Roma *et al.*, 2019).

Tras la identificación del problema a resolver, hay que evaluar el rendimiento del modelo. En tareas de clasificación se evalúa en términos de precisión (*accuracy*, en inglés) y en tareas de regresión se usa la tasa de error (*error rate*, en inglés). La precisión indica la proporción de ejemplos correctos producidos por el modelo y la tasa de error indica la proporción de ejemplos incorrectos producidos por el modelo (Goodfellow *et al.*, 2016).

Goodfellow *et al.* (2016) señala que, los algoritmos de aprendizaje automático pueden clasificarse según su experiencia durante el entrenamiento en: (a) aprendizaje supervisado, cuando se comparan las salidas que genera el modelo con las salidas ideales, se usan en tareas de clasificación y regresión (Casas Roma *et al.*, 2019; Escamilla-García *et al.*, 2020); y (b) aprendizaje no supervisado, el algoritmo es quien aprende, descubre y agrupa patrones y características de los datos de entrada, se usa en tareas de agrupamiento (Casas Roma *et al.*, 2019).

1.3.2.1. Capacidad del modelo

El reto de desarrollar una ANN es que pueda generalizar su conocimiento, es decir, que tenga un buen rendimiento con datos que nunca haya visto y no solo con datos usados en el entrenamiento. El error de entrenamiento se calcula cuando se entrena una ANN y disminuye mientras más se entrena. Sin embargo, reducirlo no es suficiente ya que solo se resolvería un problema de optimización. Se busca reducir el error de generalización, también conocido como error de prueba, con un conjunto de datos diferente al conjunto de entrenamiento (Goodfellow *et al.*, 2016).

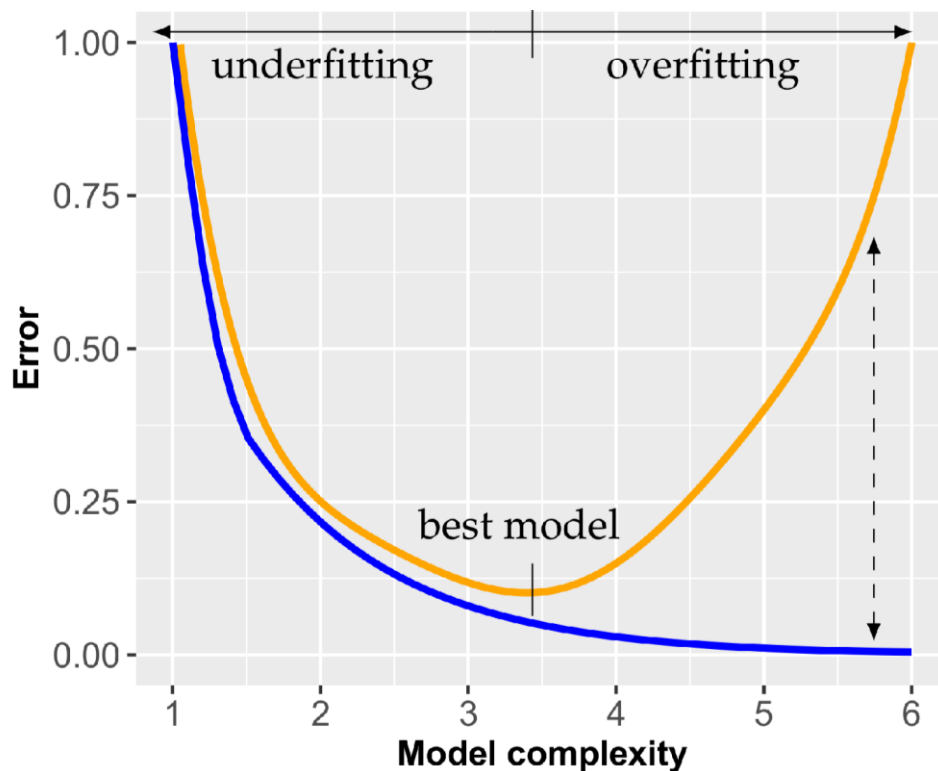
A rasgos generales, el rendimiento de un modelo se mide por: (a) su capacidad de reducir el error de entrenamiento; y (b) su capacidad para reducir la diferencia entre los errores de entrenamiento y los de prueba (Goodfellow *et al.*, 2016). Cabe mencionar que, el entrenamiento, la optimización y la generalización son conceptos que están relacionados. Además, se prevé que tanto los errores de

entrenamiento como los de prueba se reduzcan al crear modelos de ANN (François Chollet, 2017).

Una ANN puede presentar problemas durante la etapa de entrenamiento que harán que, cuando se utilicen para resolver problemas reales, no puedan generalizar su conocimiento. A estos problemas se les conoce como *underfitting* y *overfitting* y una manera de identificarlos se muestra en la Figura 4. El *overfitting* ocurre cuando una ANN conoce a la perfección el conjunto de datos. Se identifica cuando la diferencia entre los errores de entrenamiento y prueba aumenta. El *underfitting* se produce cuando una ANN no es capaz de reducir el error por más que el modelo se entrene (Goodfellow *et al.*, 2016).

Figura 4

Curvas de error idealizadas para modelos de ANN generales



Nota. Adaptado de “Evaluación de modelos, selección de modelos y error de generalización en modelos de regresión”, por Emmert-Streib y Dehmer, 2019, *Machine Learning and Knowledge Extraction*, 1(1), pág. 532. (<https://doi.org/10.3390/make1010032>). CC BY 4.0.

El problema de *underfitting* puede deberse a una falta de entrenamiento o a que el modelo es muy simple. Para el primer caso, se debe entrenar la ANN durante más épocas; para el segundo, se recomienda aumentar el número de neuronas ocultas o capas ocultas. El problema de *overfitting* se debe a que una ANN es muy compleja para la cantidad de información que se tiene (Müller & Guido, 2016). Se recomienda disminuir el número de neuronas ocultas o capas ocultas (François Chollet, 2017).

Según Goodfellow *et al.* (2016), existe diversos estudios que demuestran que la cantidad de datos utilizados para entrenar el modelo ayudan a disminuir el error de prueba. Ya que la complejidad del modelo y la variedad de datos están relacionados. De manera que, cuantos más datos se tengan, se podrá crear modelos más complejos sin llegar a presentar problemas de *overfitting* (Müller & Guido, 2016). Si bien es cierto, no existe una fórmula mágica para saber el número exacto de capas y neuronas, se debe probar diferentes configuraciones hasta encontrar la óptima para los datos (François Chollet, 2017).

1.3.2.2. Descenso del gradiente

El desarrollo de modelos de ANN consiste en maximizar o minimizar una función denominada función objetiva. Esta función se expresa, por lo general, en términos de minimización cuando se calcula el error y en términos de maximización cuando se calcula la precisión. Cuando la función objetiva es minimizada recibe el nombre de función de coste, función de pérdida o función de error. En esta investigación se referirá a ella como función de pérdida (Goodfellow *et al.*, 2016).

Según François Chollet (2017), la función de pérdida permite identificar el éxito de una ANN cuando esta realiza una tarea determinada. El autor señala que, para ello, se utilizan algoritmos de aprendizaje que determinan la manera en que una ANN actualizará sus pesos. Teniendo en cuenta la función de pérdida, el error puede calcularse como la diferencia entre lo previsto y lo obtenido por la ANN. Posteriormente, los pesos son modificados de tal manera que la ANN aprenda cual es la función de salida que se desea (Casas Roma *et al.*, 2019).

Una técnica de optimización iterativa de primer orden es el descenso de gradiente que utiliza la derivada para encontrar el mínimo de una función. Este algoritmo es uno de los más comunes para optimizar la función de pérdida de una

ANN. Asimismo, los algoritmos que usan el gradiente se denominan algoritmos de primer orden. Cabe mencionar que, existen muchos algoritmos de optimización, los cuales se mencionan más adelante (Goodfellow *et al.*, 2016).

El vector del gradiente indica, dado cualquier punto de una función, hacia donde crece la función de pérdida, es decir, hacia donde crece el error. Pero cuando se desarrollan modelos de ANN se busca reducir el error, no aumentarlo. Para lograr esto, se debería ir en la dirección opuesta del gradiente, es por esta razón, que se le conoce como el descenso del gradiente (Goodfellow *et al.*, 2016).

El Descenso del Gradiente busca encontrar un valor mínimo de una función al que se le denomina mínimo global. Este valor puede confundirse con un punto llamado mínimo local, que es un punto menor a comparación de otros puntos vecinos. Solo puede haber un mínimo global o múltiples mínimos locales y se debe evitar que la función caiga en uno de sus mínimos locales. Esto no siempre es posible, por lo que puede aceptarse un mínimo local, siempre y cuando sea un valor significativamente bajo de la función de pérdida. (Goodfellow *et al.*, 2016).

1.3.2.3. Hiperparámetros

Los hiperparámetros son ajustes que permiten controlar el comportamiento de una ANN y se establecen antes de optimizar la red (Ramachandran *et al.*, 2017). Como se ve en la Tabla 2, existen dos tipos de hiperparámetros: (a) hiperparámetros que están asociados al modelo y (b) hiperparámetros asociados con el descenso del gradiente (Vijayashanthar *et al.*, 2018). Una mala configuración de hiperparámetros puede ser, muchas veces, la causante de que una ANN no tenga capacidad de generalizar (Goodfellow *et al.*, 2016).

Tabla 2

Tipos de hiperparámetros en modelos de aprendizaje automático

Estudio	Tipo	Nombre
Vijayashanthar <i>et al.</i> (2018).	Hiperparámetros asociados al modelo.	Número de capas ocultas. Número de neuronas. Función de activación.

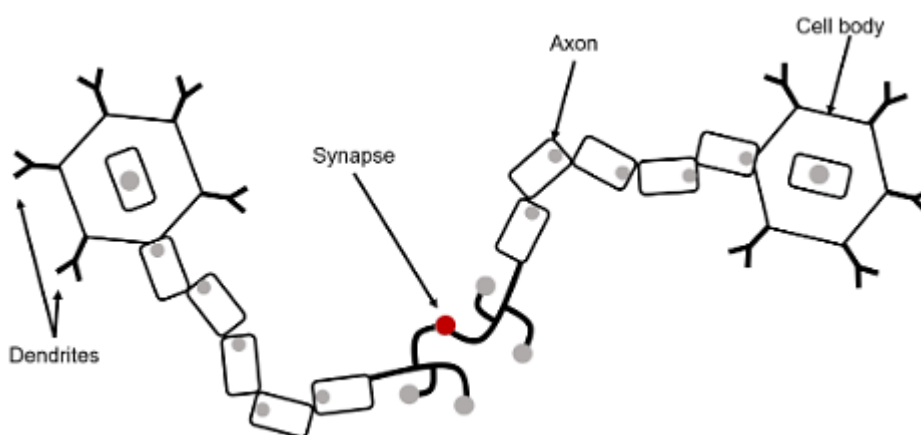
Estudio	Tipo	Nombre
	Hiperparámetros asociados al descenso del gradiente.	Tasa de aprendizaje.
		Tamaño del lote.
		Número de épocas.

1.3.3. Redes neuronales artificiales

Una ANN es un modelo informático basado en la estructura y la función de una neurona biológica. Para tener una idea, como se muestra en la Figura 5, una neurona biológica está formada por los siguientes componentes: (a) dendritas, que llevan señales eléctricas hacia (b) el cuerpo celular, las cuales añaden y recogen estas señales y (c) el axón, que recoge y lleva información del cuerpo celular a otras neuronas. Asimismo, la conexión entre un axón y la dendrita de otra neurona se conoce como sinapsis (Escamilla-García *et al.*, 2020).

Figura 5

Estructura de una neurona biológica

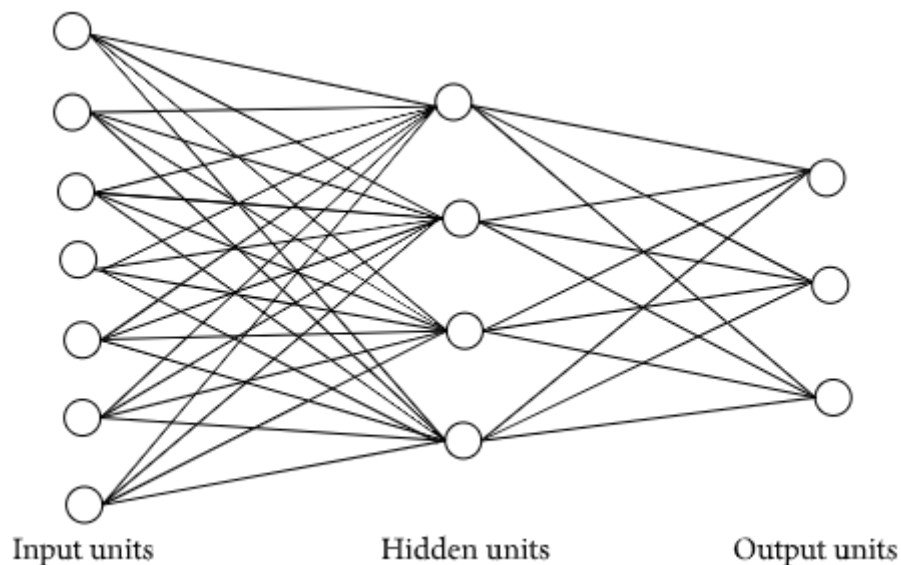


Nota. De “Aplicaciones de las redes neuronales artificiales en la tecnología de invernaderos y una visión general para el desarrollo de la agricultura inteligente” por Escamilla-García *et al.*, 2020, *Applied Sciences*, 10(11), pág. 3, (<https://doi.org/10.3390/app10113835>), CC BY.

De manera similar, una ANN está compuesta de elementos de procesamientos (neuronas) conectados entre sí con coeficientes (pesos) fijos en cada conexión (Shanmuganathan, 2016). Comprende tres componentes principales, como se indica en la Figura 6. Cabe mencionar que los datos de entrenamiento especifican lo que se debe hacer en la capa de salida y el algoritmo de aprendizaje especifica como usar las otras capas para producir el resultado deseado (Escamilla-García *et al.*, 2020).

Figura 6

Estructura de una red neuronal artificial



Nota. De “Uso de un modelo de red neuronal artificial perceptrón multicapa para predecir la sequía” por Z. Ali *et al.*, 2017, *Advances in Meteorology*, 2017, pág. 4, (<https://doi.org/10.1155/2017/5681308>), CC BY.

Las ANN comenzaron en la década de 1940, cuando McCulloch y Pitts propusieron un modelo con el fin de explicar el funcionamiento del cerebro humano. Para los años cincuenta, se aplicaron modelos basados en el concepto de perceptrón. Durante los años setenta y ochenta se produjo un estancamiento en la investigación de las ANN debido a que se demostraron varias limitaciones de los diferentes modelos existentes (Casas Roma *et al.*, 2019).

No es hasta el año 1982 cuando John Hopfield ayuda al progreso de las ANN al proponer su modelo de *Backpropagation* (BP) el cual superaba las limitaciones del perceptrón y facilitaba la fase de aprendizaje de las ANN. Posteriormente, en los años ochenta y noventa, hubo una decaída ya que los modelos no eran capaces de cumplir las expectativas, sin embargo, en las investigaciones se aplicaban otros modelos que parecían más prometedores (Casas Roma *et al.*, 2019).

En la primera década del siglo XXI, reaparecieron las ANN, cuando Geoffrey E. Hinton, junto a otros investigadores, presentaron un algoritmo que permitía entrenar a las ANN de forma eficiente. A ellos le siguieron más investigadores llegando a popularizar el término *Deep Learning*, que hacía referencia a entrenar redes más profundas (Casas Roma *et al.*, 2019).

Las ANN ayudan a resolver problemas complejos del mundo real. Su popularidad radica en como procesan la información para mejorar su aprendizaje, no linealidad y su capacidad de generalización. También se les conoce como redes neuronales prealimentadas (*feedforward neural networks*, en inglés), debido a que la salida de una determinada neurona se utiliza como entrada de la neurona siguiente (Abiodun *et al.*, 2018).

Las ANN pueden ser diseñadas y entrenadas de manera similar a otros modelos de aprendizaje automático, usando el descenso del gradiente u optimizadores similares. Por medio de un proceso de iteraciones, los optimizadores basados en el gradiente reducen la función de pérdida a valores muy bajos. La principal diferencia entre los modelos lineales y las ANN es la no linealidad, que permite tener funciones de pérdidas no convexas (Goodfellow *et al.*, 2016).

1.3.3.1. Perceptrón

El perceptrón fue creado en el año 1962 por Rosenblatt, asimismo, es la unidad más básica de las ANN y ayuda a resolver problemas sencillos y su operación es similar a la regresión no lineal (Casas Roma *et al.*, 2019). El Perceptrón está compuesto por una neurona conectada a una capa de entrada con su respectiva salida (Zhang *et al.*, 2018).

Según Wang *et al.* (2020), el perceptrón sin función de activación puede expresarse de la siguiente manera:

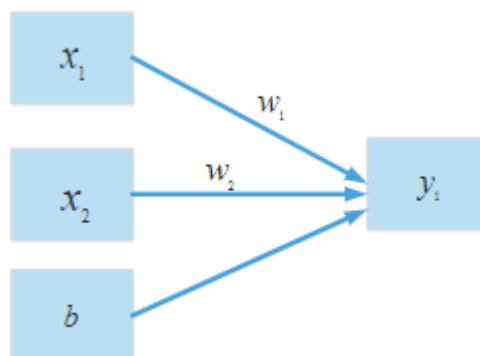
$$y_1 = w_1x_1 + w_2x_2 + b \quad (1.1)$$

$$y_1 = \sum_{i=1}^n w_ix_i + b \quad (1.2)$$

Siguiendo las ecuaciones (1.1) y (1.2), la salida del perceptrón está dada por la sumatoria de las entradas multiplicada por sus respectivos pesos, a esta cantidad se le suma el sesgo, como se ilustra en la Figura 7.

Figura 7

Perceptrón sin función de activación



Nota. De “La influencia de la función de activación en un modelo de red neuronal de convolución para el reconocimiento de expresiones faciales” por Y. Wang *et al.*, 2020, *Applied Sciences*, 10(5), pág. 3, (<http://dx.doi.org/10.3390/app10051897>). CC BY.

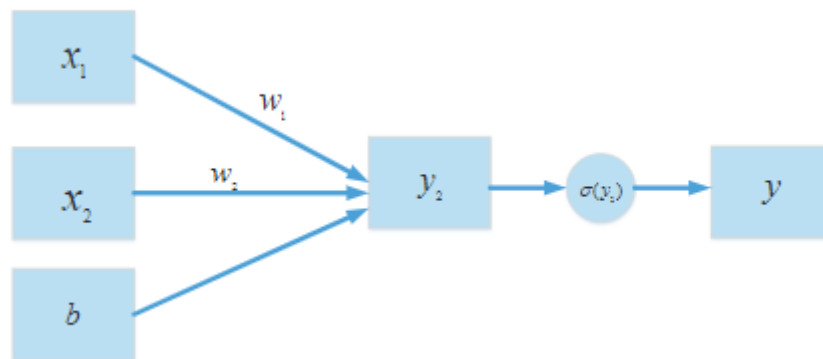
Sin embargo, Wang *et al.* (2020) señala que la definición anterior solo permite resolver problemas lineales. Debido a esto, es necesario añadir una función de activación no lineal la cual permitirá al perceptrón resolver problemas no lineales por medio de las ecuaciones (1.3) y (1.4), como se ve en la Figura 8.

$$y_2 = w_1x_1 + w_2x_2 + b \quad (1.3)$$

$$y = \sigma(y_2) \quad (1.4)$$

Figura 8

Perceptrón con función de activación



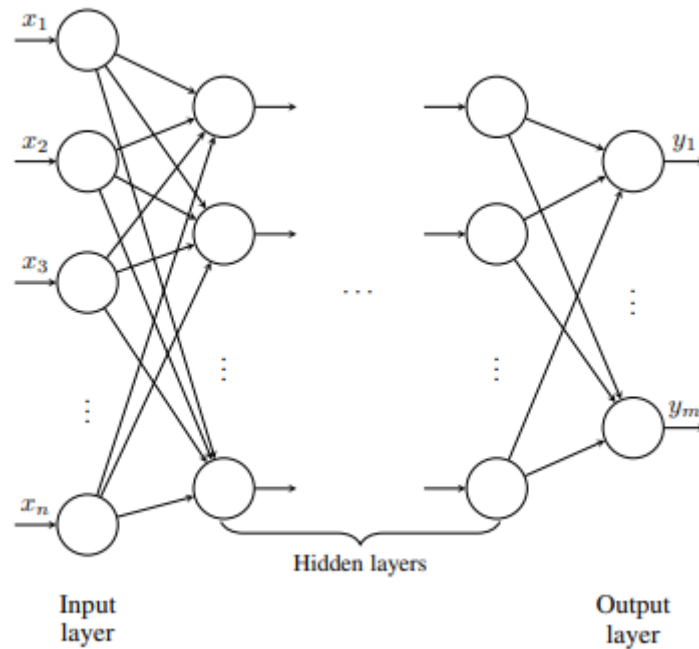
Nota. De “La influencia de la función de activación en un modelo de red neuronal de convolución para el reconocimiento de expresiones faciales” por Y. Wang *et al.*, 2020, *Applied Sciences*, 10(5), pág. 3, (<http://dx.doi.org/10.3390/app10051897>). CC BY.

1.3.3.2. Perceptrón multicapa

El perceptrón multicapa (MLP) es el tipo de ANN más usado e investigado en la actualidad (Ncibi *et al.*, 2017). Como se ve en la Figura 9, el MLP está formado por tres componentes principales (Ecer *et al.*, 2020). El MLP conecta un conjunto de entradas sobre un conjunto de salida (Saha *et al.*, 2021) que contiene múltiples capas ocultas (Jassam Mohammed *et al.*, 2020) conectadas unas tras otras (Saha *et al.*, 2021).

Figura 9

Arquitectura general de un perceptrón multicapa



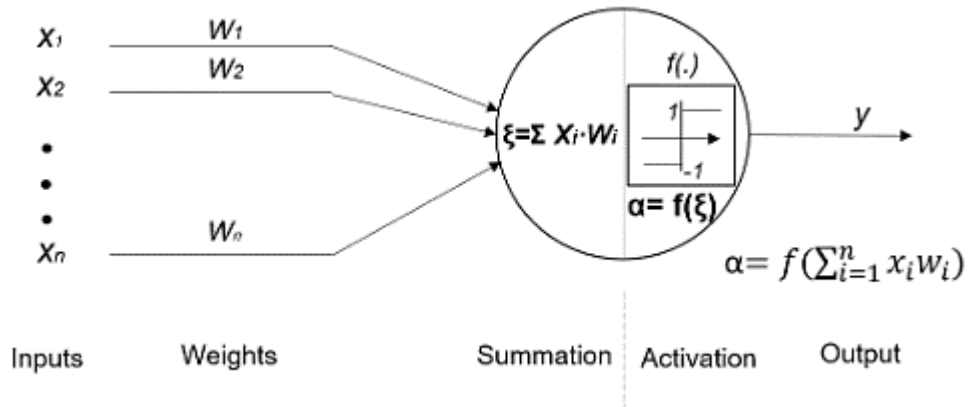
Nota. De “Una revisión sistemática de los análisis de electroencefalografía basados en el aprendizaje profundo” por Y. Roy *et al.*, (2019), *Journal of Neural Engineering*, 16(5), pág. 5, (<https://doi.org/10.1088/1741-2552/ab260c>). CC BY.

1.3.3.3. La neurona

La neurona es el componente más básico de una ANN. Es la encargada de aplicar una determinada función a los valores de entrada que recibe de otras neuronas, dando como resultado una nueva salida. Como se ve en la Figura 10, una neurona artificial consta de tres partes, al igual que una neurona biológica. Es decir, tiene un conjunto de entradas que trae información de otras neuronas, una unidad de procesamiento que suma las entradas con sus respectivos pesos antes de aplicarles una función de activación, y una salida que lleva la información a otras neuronas (Escamilla-García *et al.*, 2020).

Figura 10

Esquema básico de una neurona artificial



Nota. Una neurona tiene un conjunto de entradas denotadas como $X = \{x_1, x_2, \dots, x_n\}$, las cuales están afectadas por $W = \{w_1, w_2, \dots, w_n\}$, donde w_j^i es el peso o importancia de las neuronas. De “Aplicaciones de las redes neuronales artificiales en la tecnología de invernaderos y una visión general para el desarrollo de la agricultura inteligente” por Escamilla-García *et al.*, 2020, *Applied Sciences*, 10(11), pág. 4, (<https://doi.org/10.3390/app10113835>), CC BY.

El procesamiento que lleva la información de otras neuronas y las suma con sus pesos se realiza mediante una función llamada función de transferencia. En términos generales, esta función es la encargada de combinar cada una de las entradas con sus distintos pesos. Cabe mencionar que la función más usada es la función de suma ponderada, sin embargo, la elección de una función dependerá del problema y de los datos (Casas Roma *et al.*, 2019).

A continuación, se presentan los tipos de función de transferencia que hay (Casas Roma *et al.*, 2019).

- Función suma ponderada:

$$z(x) = \sum_{j=1}^n x_j \cdot w_j^i \quad (1.5)$$

- Función máximo:

$$z(x) = \max (x_1 w_1^i, \dots, x_n w_n^i) \quad (1.6)$$

- Función mínimo:

$$z(x) = \min (x_1 w_1^i, \dots, x_n w_n^i) \quad (1.7)$$

- Función lógica AND o OR:

$$z(x) = \min (x_1 w_1^i \wedge \dots \wedge x_n w_n^i) \quad (1.8)$$

$$z(x) = \min (x_1 w_1^i \vee \dots \vee x_n w_n^i) \quad (1.9)$$

1.3.3.4. Función de pérdida

Como ya se vio en el apartado de inteligencia artificial, en la parte del descenso del gradiente específicamente, las funciones de pérdida son un componente de las ANN que ayudan a evaluar el rendimiento de un modelo. Todas las funciones de pérdidas utilizadas para el desarrollo de modelos de ANN son las mismas que se usan para evaluar modelos lineales (Goodfellow *et al.*, 2016).

François Chollet (2017) menciona una serie de pautas que ayudan a elegir una determinada función de pérdida, estas dependen del tipo de tarea que se necesita resolver. Así pues, las funciones de pérdidas que existen y pueden utilizarse en las ANN son las siguientes:

- La entropía cruzada (*binary cross-entropy*) es una función de pérdida que se utiliza en problemas de clasificación de dos clases con una probabilidad como resultado esperado.
- La entropía cruzada categórica (*categorical cross-entropy*) es una función que se usa en problemas de clasificación con más de 2 clases.

- El error cuadrático medio (MSE) es la función de pérdida por excelencia para problemas de regresión.
- El error absoluto medio (MAE) se utiliza como una métrica durante el entrenamiento, junto con la validación cruzada, ayuda a determinar el número de épocas óptimo para el modelo.

1.3.3.5. Unidades de salida

El tipo de función de activación que utiliza la neurona para calcular la salida de la ANN se denomina tipo de unidad de salida (Szandala, 2021). Asimismo, su elección está relacionada con el tipo de función de pérdida elegida. Esto significa que la función de activación influye en cómo se representa la salida y, a su vez, determina la forma que tendrá la función de pérdida. En este punto, es importante mencionar que cualquier función de activación utilizada en la salida de una ANN también puede usarse en las capas ocultas (Goodfellow *et al.*, 2016).

Las funciones de activación son uno de los hiperparámetros que están asociados con el modelo de ANN (Vijayashanthar *et al.*, 2018). Su funcionamiento consiste en hacer que el modelo decida si una neurona es activada o no (Nwankpa *et al.*, 2018). Estas funciones tienen un efecto en el entrenamiento y desempeño del modelo, por lo que la elección correcta ayudará a que el modelo tenga un buen aprendizaje y capacidad de generalización (Ramachandran *et al.*, 2017).

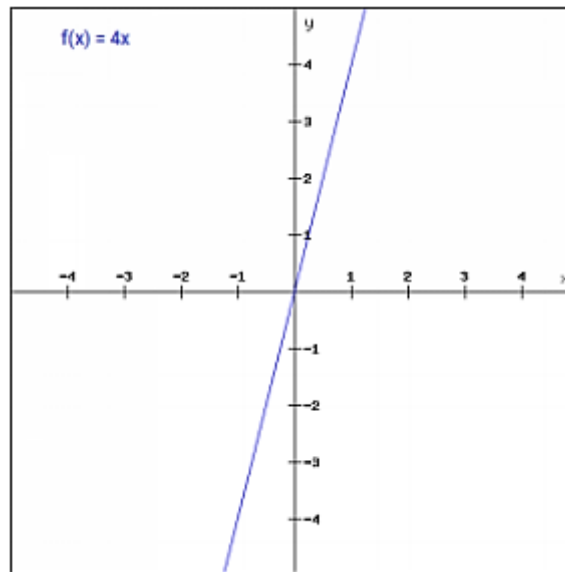
La función de activación cumple el importante papel de permitir que una ANN pueda ajustarse a curvas y superficies no lineales. Para ello, cada neurona de una ANN lo hace utilizando una función de activación no lineal sobre los datos que obtiene de otras neuronas, es decir, la suma de sus entradas. El modelo se reducirá a ser funcionalmente idéntico a una simple regresión lineal si se emplea una función de activación lineal en todo momento, incluso en ANN de varias capas (Godfrey & Gashler, 2015).

- **función de activación Lineal:** Permite hacer combinaciones lineales y predecir valores en cualquier rango (François Chollet, 2017). Como se ve en la Figura 11, esta función es directamente proporcional a su entrada (Sharma *et al.*, 2020). Aunque su gradiente no es 0, el problema es que la derivada de la función siempre es constante, haciendo que la última

neurona permanezca lineal a los valores de la neurona anterior ya que la ANN no minimiza el error. Por esta razón, son utilizadas en la capa de salida en problemas de regresión (Szandała, 2021).

Figura 11

Función de activación Lineal

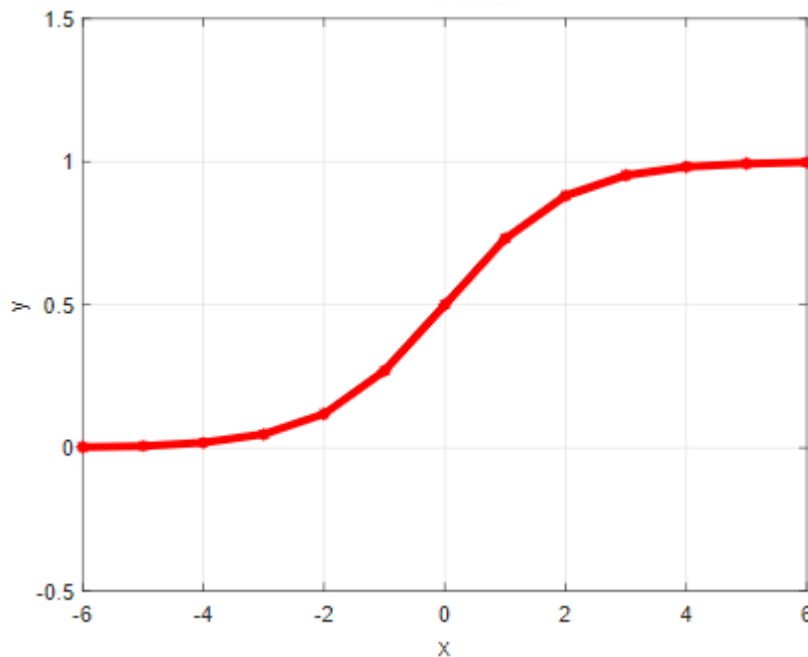


Nota. De “Activation Functions in Neural Networks” por S. Sharma *et al.*, 2020, *International Journal of Engineering Applied Sciences and Technology*, 4(12), pág. 312, (<https://doi.org/10.33564/IJEAST.2020.v04i12.054>). CC BY.

- **función de activación Sigmoide:** Función no lineal que ha sido una de las más utilizadas. La principal ventaja es que está acotada, dando resultados entre 0 y 1 (Szandała, 2021). Son aplicadas en clasificación binaria, regresión logística, entre otros (Nwankpa *et al.*, 2018). El problema de esta función es que al acercarse a los valores límites, la ANN aprende despacio o, inclusive, deja de aprender; además, no está centrada en el 0 (ver Figura 12), lo que complica el aprendizaje ya que los pesos no pueden aumentar y disminuir a la vez (Szandała, 2021).

Figura 12

Función de activación Sigmoide



Nota. De “La influencia de la función de activación en un modelo de red neuronal de convolución para el reconocimiento de expresiones faciales” por Y. Wang *et al.*, 2020, *Applied Sciences*, 10(5), pág. 5, (<http://dx.doi.org/10.3390/app10051897>). CC BY.

- **función de activación SoftMax:** Es una combinación de muchas funciones sigmoideas (Nwankpa *et al.*, 2018). Al igual que una función sigmoidea da valores entre 0 y 1. Estos pueden ser tratados como probabilidades para una clase en particular. Para los problemas de clasificación multiclase, se utiliza la función SoftMax en la capa de salida como función de activación (Sharma *et al.*, 2020). Es decir, cuando hay más de 2 clases y la suma de las puntuaciones es 1 (François Chollet, 2017, pág. 77).

1.3.3.6. Unidades ocultas

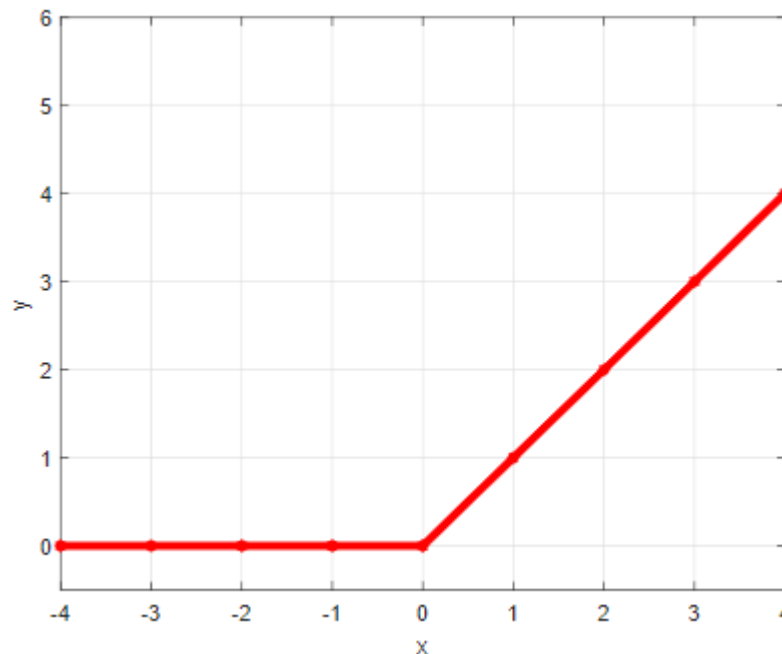
Las funciones de las neuronas ocultas toman el valor de la función de entrada y lo modifican para que cada neurona haga más que aproximaciones lineales (Casas

Roma *et al.*, 2019). Aunque, es complicado saber qué tipo de función de activación hará que el modelo se desempeñe de manera óptima, el proceso consiste en un ensayo de prueba y error: (a) intuir que función de activación puede hacer que la neurona oculta funcione bien, (b) entrenar el modelo con la función de activación elegida, y (c) evaluar el rendimiento del modelo (Goodfellow *et al.*, 2016).

- **función de activación ReLU:** Las *Rectified Linear Units* fueron propuestas por Nair y Hinton en 2010 (Nwankpa *et al.*, 2018). Son la principal elección, las más usadas y exitosas (Ramachandran *et al.*, 2017). Son fáciles de optimizar al igual que una función lineal (Goodfellow *et al.*, 2016). Como se ve en la Figura 13, pone los valores negativos en 0 (Szandała, 2021). Aunque a veces esto puede ser un problema debido a que los pesos no se actualizan, dificultando el aprendizaje ya que las neuronas no vuelven a aprender (Goodfellow *et al.*, 2016).

Figura 13

Función de activación ReLU

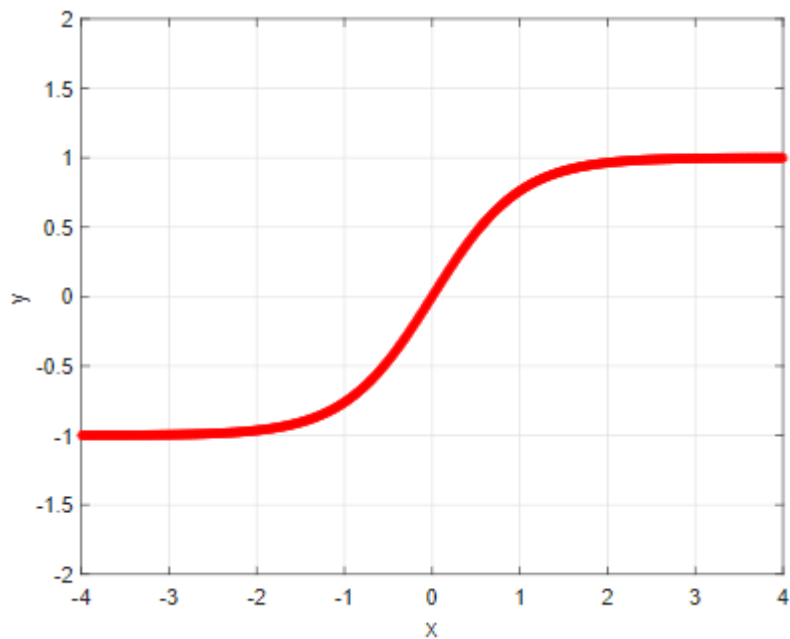


Nota. De “La influencia de la función de activación en un modelo de red neuronal de convolución para el reconocimiento de expresiones faciales” por Y. Wang *et al.*, 2020, *Applied Sciences*, 10(5), pág. 5, (<http://dx.doi.org/10.3390/app10051897>). CC BY.

- **función tangente hiperbólica:** Sus resultados están entre -1 y 1, es decir, entradas negativas tendrán salidas negativas (Szandała, 2021). Una de las ventajas es que produce salidas centradas en 0, como se ve en la Figura 14 (Nwankpa *et al.*, 2018). Al igual que la Sigmoide, tiene el problema del Desvanecimiento del Gradiente donde, conforme se entrena y aplica el *backpropagation*, llega un momento que el error propagado a las capas anteriores toma un valor cada vez más pequeño que prácticamente no se corrige nada y la ANN no aprende (Goodfellow *et al.*, 2016).

Figura 14

Función de activación Tangente hiperbólica



Nota. De “La influencia de la función de activación en un modelo de red neuronal de convolución para el reconocimiento de expresiones faciales” por Y. Wang *et al.*, 2020, *Applied Sciences*, 10(5), pág. 5, (<http://dx.doi.org/10.3390/app10051897>). CC BY.

1.3.3.7. Método Backpropagation

Escamilla-García *et al.* (2020) señala que el *backpropagation* (BP) es un tipo de entrenamiento de ANN que se emplea en la aplicación del aprendizaje supervisado. El BP implica propagar proporcionalmente el error de la ANN a cada una de las neuronas de entrada que han influenciado en el error final (Casas Roma *et al.*, 2019). En primer lugar, se realiza una propagación hacia adelante (*feedforward*) que consiste en introducir la instancia de entrenamiento para luego obtener la salida de la red. En segundo lugar, se determina el error cometido por la capa de salida, luego este error es propagado hacia atrás para hallar la función que permita modificar los pesos correspondientes (Casas Roma *et al.*, 2019).

Saikia *et al.* (2020) indica una serie de pasos para entender el funcionamiento del BP:

1. Los pesos de la red se inicializan aleatoriamente.
2. El conjunto de entrenamiento es presentado a la red.
3. La entrada es propagada hacia adelante para generar la salida.
4. Se calcula el error comparando los valores esperados y los predichos.
5. El error es propagado hacia atrás a toda la red.
6. Los pesos son ajustados para minimizar el error.
7. Se repiten los pasos del 2 al 6 y el proceso se repite hasta que el error no disminuya más.

1.3.4. Desarrollo de los modelos

La cantidad de investigaciones sobre el uso de las ANN ha aumentado, sin embargo, la mayoría de trabajos no proporcionan una evaluación del rendimiento ni una justificación de las elecciones metodológicas en el desarrollo de los modelos. Es importante conocer los factores que afectan al rendimiento como, por ejemplo, las entradas adecuadas para la ANN, los datos y su procesamiento, la arquitectura y parámetros que ayudan a la optimización y validación del modelo (Lathuiliere *et al.*, 2020).

1.3.4.1. División del conjunto de datos

Los datos recopilados deben utilizarse para crear los conjuntos de entrenamiento, validación y prueba (François Chollet, 2017). El conjunto de prueba suele ser el 20 % del conjunto de datos (SINGH & SINGH, 2016). De los datos restantes, el 80 % será para el entrenamiento y el 20 % para la validación (Goodfellow *et al.*, 2016). Durante el entrenamiento, se utilizan tanto el conjunto de entrenamiento como el de validación; sin embargo, el de validación se utiliza para ajustar el modelo. La ANN se evalúa con el conjunto de prueba una vez ajustado el modelo (François Chollet, 2017).

Un conjunto de datos que ha sido utilizado para entrenar un modelo de ANN no puede utilizarse para evaluar su rendimiento, esto se debe a la filtración de información de los datos (François Chollet, 2017). Cuando se ajusta un parámetro puede filtrarse muy pocos bits de información, por lo que la ANN seguirá siendo fiable. Sin embargo, son muchos los parámetros y configuraciones a probar, a consecuencia, se filtrará muchos bits de información de los datos y la ANN no funcionará bien para datos nuevos, puesto que la ANN recordará los datos y el error minimizado será erróneo (Müller & Guido, 2016).

Algunos puntos a tener en cuenta a la hora de dividir los datos según François Chollet (2017) son:

- a. la representatividad de los datos,
- b. la flecha del tiempo, y
- c. la redundancia de los datos.

Para tener una buena representación de los datos se recomienda mezclar los datos al azar y luego dividirlos. En tareas de clasificación, cuando no se mezclan los datos, suele suceder que el conjunto de entrenamiento contiene solo las clases de 0 al 5, por ejemplo, y el conjunto de validación las clases 6 y 7. Para ANN que tratan de predecir el futuro como, por ejemplo, el clima para mañana. Para este caso, no se debe mezclar los datos de manera aleatoria, sino que se debe buscar que los datos del conjunto de validación tengan una fecha posterior que los datos del conjunto de entrenamiento (François Chollet, 2017).

Existe redundancia en los datos cuando algunos ejemplos aparecen en el conjunto de entrenamiento y también en el conjunto de validación. Cuando sucede esto, se está evaluando el modelo con datos que han sido utilizados durante la

etapa de entrenamiento. Este problema suele ocurrir cuando se trabaja con datos del mundo real y la mejor solución es asegurarse que el conjunto de entrenamiento y validación no tenga datos iguales (François Chollet, 2017).

1.3.4.2. Validación cruzada

Cuando se tiene muy pocos ejemplos es difícil dividir el conjunto de datos en 3 partes. En primer lugar, porque el conjunto de validación sería muy pequeño, por ejemplo, 100 datos. En segundo lugar, podría existir una gran variabilidad en torno al error estimado cuando se ajusta el modelo, por lo que se haría difícil determinar si un modelo es mejor que otro. Estos problemas no ocurren cuando se entrena un modelo con un conjunto de datos que tiene ciento de miles de ejemplos (Goodfellow *et al.*, 2016).

La validación cruzada es una técnica que se aplica como solución a los problemas ya mencionados. El procedimiento consiste en fraccionar el conjunto de entrenamiento en K secciones de igual tamaño de forma aleatoria (Goodfellow *et al.*, 2016). El valor de K puede ser 5 o 10. Donde, para cada partición i , se entrena el modelo con las particiones restantes $K - 1$, y se evalúa con la partición i (Müller & Guido, 2016). Al finalizar, la puntuación del modelo será el promedio de las K puntuaciones obtenidas (François Chollet, 2017). Cabe mencionar que, el modelo se divide y entrena repetidas veces, determinadas por K (Müller & Guido, 2016).

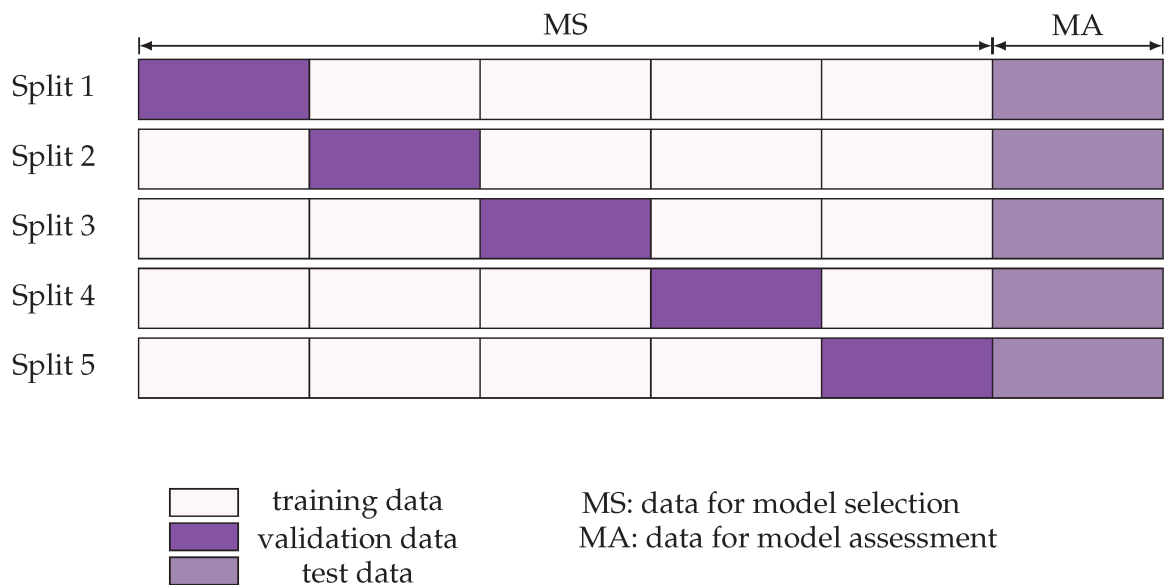
Para entender mejor el procedimiento, en la Figura 15 se muestra un ejemplo donde se ha realizado una validación cruzada con $K = 5$. El conjunto de datos se divide en dos partes: un conjunto de entrenamiento y un conjunto de prueba. Luego, el conjunto de entrenamiento es dividido en $K = 5$ partes de, aproximadamente, igual tamaño; donde $K - 1 = 4$ son utilizadas para el entrenamiento y 1 para la validación. A continuación, se entrena una secuencia de modelos utilizando el conjunto de entrenamiento (MS):

1. El primer modelo (*Split 1*) se entrena usando la partición 1 como conjunto de validación y las particiones restantes (2, 3, 4 y 5) como conjunto de entrenamiento.
2. Se construye el segundo modelo (*Split 2*), esta vez usando la partición 2 como conjunto de validación y las particiones 1, 3, 4 y 5 como conjunto de entrenamiento.

3. El proceso se repite para el tercer modelo (*Split 3*), se usa la partición 3 como conjunto de validación y las particiones 1, 2, 4 y 5 para el entrenamiento.
4. El procedimiento se repite para el cuarto y quinto modelo usando sus respectivas particiones.
5. Para cada uno de los modelos se calcula el error. Para este ejemplo se tendrán 5 resultados.
6. Finalmente, se calcula el error promedio.
7. Luego de haber realizado la validación cruzada se vuelve a entrenar el modelo con todo el set de entrenamiento y se evalúa con el set de prueba.

Figura 15

Visualización de la validación cruzada



Nota. Adaptado de “Evaluación de modelos, selección de modelos y error de generalización en modelos de regresión”, por Emmert-Streib y Dehmer, 2019, *Machine Learning and Knowledge Extraction*, 1(1), pág. 541. (<https://doi.org/10.3390/make1010032>). CC BY 4.0.

1.3.4.3. Preprocesamiento de los datos

Para un correcto preprocesamiento de los datos, una ANN requiere que los datos de entrada tengan unas características específicas (Casas Roma *et al.*, 2019). Por ejemplo, cada dato utilizado debe ser independiente entre sí, y los conjuntos de entrenamiento y prueba deben estar distribuidos de forma similar (Goodfellow *et al.*, 2016). Según François Chollet (2017), el proceso de aprendizaje de los modelos se puede facilitar haciendo que los datos estén entre 0 y 1 y estos sean homogéneos.

Según Casas Roma *et al.* (2019), si un conjunto de datos tiene atributos categóricos, estos deben convertirse en atributos numéricos. Asimismo, estos atributos deben tener un rango comprendido entre 0 y 1; para ello, se realiza una estandarización de los datos. Müller & Guido (2016) señala que, la estandarización de los datos puede hacerse de diferentes, las cuales se describen a continuación:

- *Standard scaler* hace que cada dato tenga una media igual a 0 y una varianza igual a 1. Sin embargo, este no asegura un valor mínimo ni máximo para los datos.
- *Robust scaler* es similar al *Standard Scaler*, sin embargo, usa la mediana y los cuartiles en lugar de la media y la varianza. El problema es que puede ignorar valores muy diferentes del resto.
- *Min max scaler* encuentra los valores mínimos y máximos de un conjunto de datos, y luego escala los datos a valores comprendidos entre 0 y 1. Es uno de los más utilizados.
- *Normalizer* escala cada dato con la inversa de su longitud en un círculo con un radio de 1. Solo se usa cuando la dirección de los datos es más importante que su longitud.

1.3.4.4. Arquitectura

La arquitectura de los modelos de ANN es una consideración esencial en su desarrollo. Es decir, determinar el número de neuronas que tendrá y cómo estas se conectarán las unas a las otras (Goodfellow *et al.*, 2016). La arquitectura o tipología de una ANN depende de cómo esté organizada su estructura y del problema que se desea resolver, asimismo, se debe tener cuidado con los problemas de *overfitting* y *underfitting* (Casas Roma *et al.*, 2019).

En primer lugar, para determinar la dimensión de la capa de entrada, Velasco *et al.* (2019) señala que la cantidad de neuronas de entrada debe ser igual que el número de características del conjunto de entrenamiento. Además, señala que la revisión de la literatura y correlación entre variables ayudan a identificar que variables afectan al valor que se quiere predecir. Por el contrario, Casas Roma *et al.* (2019) señala que el número de entradas puede ser establecido por el investigador.

En segundo lugar, está la dimensión de la capa de salida. El número de neuronas de salida vendrá determinado por el problema y la finalidad de la ANN. Por ejemplo, para problemas de clasificación el número de salidas se definirá a partir del número de clases. Para modelos de regresión es mucho más simple, puesto que solo pueden tener un único valor de salida (Casas Roma *et al.*, 2019).

En cuanto a la dimensión de la capa oculta, es posible que una capa oculta sea suficiente para que una ANN aproxime cualquier función no lineal (Goodfellow *et al.*, 2016; Velasco *et al.*, 2018, 2015). Según Casas Roma *et al.* (2019), el número de neuronas ocultas es de 1 a 2 veces la cantidad de neuronas de entrada y, según Velasco *et al.* (2019), el número máximo sin que se produzca overfitting está dado por la ecuación (1.10) de Stathakis’.

$$Nh = \frac{Ns}{(\alpha * (Ni + No))} \quad (1.10)$$

Donde el número de neuronas ocultas Nh se calcula dividiendo el número de datos Ns con el factor de escala arbitraria α , multiplicado por la sumatoria de la cantidad de neuronas de entrada Ni y la cantidad de neuronas de salida No . Asimismo, el factor α varía de 2 hasta 10.

1.3.4.5. Algoritmos con tasas de aprendizaje adaptativas

El ajuste de los parámetros de una ANN se realiza después de haber entrenado el modelo, cabe mencionar que, cuando se habla de optimización y ajuste de un modelo se está refiriendo a expresiones sinónimas. Para ello, los optimizadores son los encargados de encontrar los parámetros (pesos) de una función de pérdida. Asimismo, esta permite cuantificar la calidad de los pesos (Lathuiliere *et al.*, 2020).

Establecer la tasa de aprendizaje (*learning rate*, en inglés) es importante para optimizar un modelo de ANN. Puede llegar a ser un paso problemático puesto que, una tasa de aprendizaje demasiado grande hará que la ANN no alcance el mínimo deseado; mientras que, si se establece una tasa de aprendizaje demasiado pequeña hará que el sistema sea lento para converger. En la práctica, este valor se disminuye gradualmente (Lathuiliere *et al.*, 2020; Pack *et al.*, 2016).

Existen distintos algoritmos utilizados para entrenar los modelos de ANN, por ejemplo, aquellos que no usan la derivada (Mesejo *et al.*, 2015). Ahora bien, por las razones mencionadas, se han propuesto muchas mejoras del algoritmo del descenso del gradiente básico. Asimismo, se han desarrollado diversos algoritmos con tasas de aprendizaje adaptativas. Estos son algunos de los métodos de optimización más utilizados en la actualidad (Lathuiliere *et al.*, 2020).

- **AdaGrad (*Adaptive Gradient*)**: Es un optimizador que fue propuesto por primera vez en 2010 y 2011 y ha crecido en popularidad en los últimos años (Ward *et al.*, 2018). Según Lathuiliere *et al.* (2020), el método ajusta la tasa de aprendizaje de cada peso dividiéndola por la raíz cuadrada de sus valores históricos al cuadrado. La disminución de la tasa de aprendizaje es rápida para los pesos con gradiente altos, mientras que los pesos con actualizaciones pequeñas tendrán una disminución pequeña de su tasa de aprendizaje.
- **RMSProp (*Root Mean Squared Propagation*)**: Es un optimizador que fue propuesto por Geoffrey Hinton (Yixiang Wang *et al.*, 2019). Modifica el algoritmo AdaGrad con la finalidad de evitar que la tasa de aprendizaje se reduzca de forma agresiva (Lathuiliere *et al.*, 2020). Mantiene un promedio móvil del gradiente cuadrado para cada peso y divide el gradiente por la raíz cuadrada de este valor (S. Roy *et al.*, 2017). Ha demostrado buen rendimiento ya que acelera bastante la convergencia en el entrenamiento de ANN (Carlson *et al.*, 2015; Dauphin *et al.*, 2015).
- **AdaDelta**: Es un algoritmo propuesto por Zeiler. A diferencia de los otros algoritmos, este método da a cada uno de los parámetros una tasa de aprendizaje diferente. Además, para los parámetros frecuentemente actualizados se realizan cambios pequeños y para los parámetros que no se

actualizan con frecuencia los cambios son grandes (Yixiang Wang *et al.*, 2019).

- **Adam (*Adaptative Moment*):** Este algoritmo fue propuesto por Kingma *et al.* en el año 2015. El optimizador Adam combina las ventajas que ofrece el algoritmo de optimización del momento y el algoritmo AdaGrad con la finalidad de acelerar la convergencia en las primeras fases del entrenamiento (Yixiang Wang *et al.*, 2019). Según Yi *et al.* (2020), es el algoritmo más utilizado en la actualidad.

1.3.4.6. Número de epochs

Se habla de una época (*epoch*, en inglés) cuando todo el conjunto de entrenamiento pasan hacia adelante y hacia atrás por la ANN (Y. Roy *et al.*, 2019). Entrenar una ANN en una sola época no es buena idea ya que la ANN no aprende lo suficiente. Por esta razón, la actualización de los pesos debe hacerse en varias épocas. También es importante mencionar que conforme aumenta el número de épocas, la ANN va pasando de un ajuste insuficiente a un ajuste satisfactorio hasta un ajuste excesivo (Amirabadi *et al.*, 2020).

El número de épocas es diferente para cada modelos de ANN, aunque, está relacionado con la diversidad de los datos (Amirabadi *et al.*, 2020). Para medir la influencia de esto en el rendimiento, se debe trabajar con un número definido de épocas durante el proceso de optimización de los hiperparámetros (Diaz *et al.*, 2017; Kim, 2019). Si un modelo aprende con demasiada lentitud, hay que aumentar el número de épocas (YILDIRIM, 2019).

El número de épocas inicial se puede elegir después de hacer unas pruebas preliminares. Se debe encontrar un equilibrio entre la velocidad del entrenamiento y la precisión en el conjunto de validación. Posteriormente, se elige la mejor configuración de hiperparámetros que será evaluado con el conjunto de pruebas (Diaz *et al.*, 2017). Generalmente se utiliza una escala logarítmica para investigar el número de épocas puesto que los efectos que se tienen son multiplicativos, por ejemplo en rangos como, {10, 30, 100, 300, 800, 1000, 1500} (Dinamarca, 2018).

Una manera de determinar el número óptimo de épocas es mediante la validación cruzada. Para ello, se entrena el modelo con un número grande de épocas y un tamaño de lote con un valor de 1 para ver como varía el error, en este

caso se usa el MAE como métrica. El número de épocas donde el error de validación comienza a aumentar se determina como el más óptimo para el modelo; puesto que, pasado ese número se produciría *overfitting* (François Chollet, 2017).

1.3.4.7. Selección del batch size

Según Goodfellow *et al.* (2016), el tamaño de lote (*batch size*, en inglés) es el número de datos de entrenamiento utilizados en cada lote, a diferencia de una iteración que es el número de lotes que se necesitan para completar una época. En función al número de datos utilizados en un lote, se clasifica en: (a) descenso del gradiente por lotes, cuando se usa todo el conjunto de datos de entrenamientos; (b) descenso del gradiente por mini lotes, que usa una cantidad determinada de datos; y (c) descenso de gradiente estocástico, el cual usa un solo dato en cada iteración.

Lo ideal es utilizar todos los datos en cada iteración, sin embargo, cuando se tiene limitaciones de hardware es recomendable establecer un tamaño de lote máximo (Balles *et al.*, 2017). Asimismo, si se tiene una gran cantidad de datos para el entrenamiento, se debe dividirse en lotes más pequeños (Amirabadi *et al.*, 2020). Esto se debe a que la memoria de la GPU limita el número de datos que se pueden procesar a la vez. Los valores usados son potencia de dos, por ejemplo, un tamaño de lote de 2, 4, 8, 16, 32, 64, 128 y 256 (Balles *et al.*, 2017).

1.3.4.8. Optimización Bayesiana

La optimización bayesiana ayuda a elegir de forma eficiente el valor de los hiperparámetros. En primer lugar, se prueba una serie de hiperparámetros de forma aleatoria y de acuerdo a los resultados que se han tenido, va eligiendo el siguiente valor de hiperparámetros a probar. La decisión de cuál será el nuevo valor a probar se basa en la relación entre exploración y explotación. Con la exploración se prueban nuevos valores en los que se podría tener una mejora del modelo y con la explotación se busca nuevos valores muy cercanos a los que han dado buenos resultados (Koo *et al.*, 2018).

La optimización Bayesiana es una de las técnicas más potentes a la hora de encontrar valores objetivos. Lo que hace eficaz a esta técnica es que incorpora la creencia previa sobre el problema, lo que ayuda a dirigir el muestreo. Esto

compensa la exploración y explotación del espacio de búsqueda. Se llama Bayesiana porque utiliza el teorema de Bayes. El resultado previo representa la creencia que se tiene sobre el espacio de posibles hiperparámetros objetivos, por lo que se puede suponer que se conocen alguna de sus propiedades haciendo que algunos valores posibles sean más recomendables que otros (Brochu *et al.*, 2010).

1.4. Formulación del problema

¿Cuán preciso es el desarrollo de modelos de redes neuronales artificiales para predecir las propiedades mecánicas del suelo?

1.5. Justificación e importancia del estudio

Varios autores han investigado la creación de modelos ANN para predecir las propiedades mecánicas del suelo. Entre las principales ventajas encontradas al aplicar modelos de ANN está su capacidad de representar complicadas vinculaciones entre las variables de entrada y salida, así como, su posibilidad de aprender y generalizar su conocimiento. Estas características han hecho que se acepten como herramientas fiables, pues permiten modelar el complejo comportamiento y variabilidad de los materiales de ingeniería geotécnica.

Si bien es cierto que, en otros países se han hecho diversas investigaciones sobre la aplicación de las ANN, son pocas las disciplinas de la ingeniería civil en Perú en los que se han utilizado. Por esta razón, es importante dar a conocer a la comunidad científica que incorporar el uso de modelos de ANN en ramas de ingeniería civil son de gran ayuda. Debido a que permiten ahorrar mucho tiempo a la hora de abordar problemas complicados y permiten predecir valores de manera precisa.

La presente investigación se realiza pensando en el medio ambiente puesto que, una vez desarrollado el modelo de ANN, se podrá predecir las propiedades sin necesidad de un gran número de muestras. De igual manera, esta investigación tiene en cuenta el aspecto económico porque los resultados de las propiedades del

suelo pueden obtenerse sin necesidad de realizar pruebas de laboratorio. Esto significaría un ahorro a la hora de realizar proyectos de carreteras.

1.6. Hipótesis

La precisión del desarrollo de modelos de redes neuronales artificiales para predecir las propiedades mecánicas del suelo es del 80 %.

1.7. Objetivos

1.7.1. Objetivo general

Desarrollar un modelo de redes neuronales artificiales para predecir las propiedades mecánicas del suelo en la región Lambayeque.

1.7.2. Objetivos específicos

1. Definir qué propiedades mecánicas del suelo son las que se podrán predecir con el desarrollo de modelos de ANN.
2. Determinar los hiperparámetros óptimos del modelo de ANN mediante el lenguaje de programación de Python.
3. Demostrar que los modelos de ANN hayan aprendido correctamente comparando los valores reales con los valores predichos por el modelo.

II. MATERIAL Y MÉTODO

En este capítulo se describen los materiales y métodos utilizados para alcanzar los objetivos del estudio. El tipo y el diseño de la investigación se describen en la sección 2.1. En la sección 2.2 se detallan la población y la muestra, además de los criterios de inclusión y exclusión. En la parte 2.3 se repasan las variables de la investigación. En la sección 2.4 se describen los procedimientos e instrumentos utilizados para la recopilación de datos, así como el procedimiento de recogida de datos. En la sección 2.5 se explica cómo se han analizado los datos y los procedimientos empleados, en la sección 2.6, se repasan las normas éticas; y, finalmente, la sección 2.7, los criterios de rigor científico.

2.1. Tipo y diseño de investigación

Hernández-Sampieri y Mendoza Torres (2018) indica que el tipo de investigación no solo se debe ver como una clasificación, sino como el alcance que puede llegar a tener un proceso de investigación. Esto quiere decir que el tipo de investigación engloba todo un suceso de causalidades, de las cuales dependerá la estrategia que tome el investigador para alcanzar los objetivos de su investigación. Todo esto, claro, luego de haber definido el problema a investigar.

2.1.1. Tipo de investigación

El alcance que puede tener una investigación se clasifica de diversas maneras según los autores. Sin embargo, en este estudio se hablará de la investigación propositiva. La investigación propositiva o proyectiva, según Hurtado de Barrera (2015), consiste en proponer soluciones a un determinado problema. Esto implica definir las características de los fenómenos, conceptos o variables en un determinado contexto. También la recopilación de datos sobre los conceptos en estudio, que se denominan variables.

En consecuencia, el presente investigación es de tipo proyectiva, ya que se busca desarrollar modelos de ANN que serán propuestos para predecir las

propiedades del suelo. Fue necesario especificar las variables que podrían utilizarse como entradas para predecir una característica mecánica del suelo con el fin de crear un modelo de ANN. A continuación, había que recopilar los datos adecuados de las variables ya definidas; esto se hizo para crear la base de datos que se utilizó para entrenar los modelos de ANN (Hernández-Sampieri & Mendoza Torres, 2018).

2.1.2. Diseño de investigación

Para las investigaciones se puede usar un diseño transversal o longitudinal. Cuando se trata de un diseño transversal, los datos se recogen en un único momento en el tiempo. Un estudio de diseño longitudinal, en cambio, recogerá datos en múltiples momentos para examinar cómo cambian los resultados a lo largo del tiempo (Hernández-Sampieri & Mendoza Torres, 2018). Dado que los datos se recogieron en un momento determinado, el presente estudio tuvo un diseño transversal que tuvo en cuenta las ideas mencionadas. La recogida de ejemplos para la base de datos, la búsqueda de hiperparámetros y la predicción de atributos forman parte del proceso.

2.2. Población y muestra

Limitar la población de estudio es un paso crucial. Permite identificar en qué o quienes se centrará el estudio. Asimismo, es crucial definir el tamaño adecuado de la muestra y tener claro los procedimientos que se usarán para seleccionar la muestra. Haciendo bien estos pasos y, además, teniendo en cuenta que esta parte de la investigación está ligada al tipo de estudio se podrá lograr los objetivos propuestos (Hernández-Sampieri & Mendoza Torres, 2018).

2.2.1. Población

La población se emplea para un grupo a los que se pretende aplicar los resultados de un proyecto de estudio. Por esa razón, es importante describir las características de la población de estudio para que quede bien delimitada (Hernández-Sampieri &

Mendoza Torres, 2018). También se le denomina población al conjunto de unidades a las cuales se quiere estudiar. Dejar claro qué o por quién estará conformada la población es una parte esencial de las investigaciones (Perez *et al.*, 2020).

Teniendo en cuenta los conceptos ya mencionados, la población estuvo compuesta por 285 muestras de suelo de la región Lambayeque. Las muestras de suelo se obtuvieron de los expedientes técnicos de obras de pavimentación que fueron realizadas en las provincias de Chiclayo, Ferreñafe y Lambayeque. Asimismo, toda la información reunida fue descargada del Sistema Electrónico de Contrataciones del Estado (SEACE).

En la Tabla 3 se puede ver el número de expedientes, muestras de suelo y ensayos realizados por provincia. Se tuvo acceso a 1 expediente técnico de un proyecto que entró a licitación en el año 2021, 26 expedientes técnicos de proyectos que entraron en licitación en el año 2020 y 9 expedientes técnicos de proyectos que entraron en licitación en el año 2019. El número total de expedientes técnicos descargados fue de 36.

Tabla 3

Resumen de la información recopilada del SEACE

Lugar	Expedientes	Muestras	Ensayos
Chiclayo	21	179	1074
Ferreñafe	4	35	210
Lambayeque	11	71	426
Total	36	285	1710

2.2.2. Muestra

Debido a las limitaciones de tiempo y recursos, la muestra del estudio suele ser un subconjunto de la población. Para seleccionar la muestra se requiere, en primer lugar, haber delimitado bien la población de estudio (Hernández-Sampieri & Mendoza Torres, 2018). Por otro lado, el muestreo es la selección de un

subconjunto de unidades de análisis, el cual forma parte de la población (Perez *et al.*, 2020).

En este estudio se utilizaron un total de 285 muestras de suelo de las provincias de Chiclayo, Ferreñafe y Lambayeque. Como se puede ver, todas las muestras de suelo de la población de estudio serán usadas para desarrollar los modelos de ANN. Se tomó esa decisión puesto que la cantidad de datos no es muy grande y, además, el número de datos que se utilizaba para entrenar un modelo de ANN afecta en la precisión.

Los criterios de inclusión utilizados para seleccionar la población de estudio fueron los siguientes:

- Muestras de suelo con resultados de contenido de humedad.
- Muestras de suelo con resultados de contenido de sales.
- Muestras de suelo con resultados de análisis granulométrico.
- Muestras de suelo con resultados de los límites de Atterberg.
- Muestras de suelo con resultados de clasificación USCS y AASHTO.
- Muestras de suelo con resultados de compactación.
- Muestras de suelo con resultados del ensayo de CBR.

Por otro lado, los criterios de exclusión que se utilizaron para restringir la población de estudio son descritos a continuación:

- Muestras de suelo de expedientes técnicos que no son de carreteras.
- Muestras de suelo de expedientes técnicos del año 2014 hacia atrás.
- Muestras de suelo con datos faltantes.

2.3. Variables y operacionalización

Según Hernández Sampieri *et al.* (2014), las variables de estudio son propiedades que tienen como característica la variación, la cual puede ser observada y medida. Según Cruz del Castillo y Socorro Olivares (2014), las variables pueden ser: (a) dependientes, aquellas que dependen de lo que modifica el investigador y, además, reflejan los resultados de un estudio, y (b) independientes, las cuales son controladas por el investigador y permiten probar el efecto sobre algún resultado.

2.3.1. Variables

En primer lugar, en la presente investigación se consideró como variable independiente: (a) el Modelo de ANN; y como variable dependiente: (b) la predicción de las propiedades mecánicas del suelo. La descripción conceptual para cada una de estas variables es detalla a continuación:

1. *Modelo de ANN*, los cuales son modelos computacionales que han sido basados en el funcionamiento del cerebro humano. Asimismo, estos modelos cuentan con funciones matemáticas que hacen que las predicciones sean precisas (Saffariha *et al.*, 2021).
2. *Predicción de las propiedades mecánicas del suelo*, denota la evaluación de las cualidades mecánicas del suelo según sus propiedades físicas. Esta variable es importante ya que se podrá estimar el comportamiento que tendrá el suelo según sus propiedades (Özyankı, 2019).

2.3.2. Operacionalización

El proceso de convertir una variable conceptual en una variable operativa se conoce como operacionalización de la variable (Hernández-Sampieri & Mendoza Torres, 2018). La variable independiente del presente estudio se dividió en tres dimensiones que son: el conjunto de datos, los hiperparámetros asociados al modelo y los hiperparámetros asociados al descenso del gradiente; como se muestra en la Tabla 4. A su vez, estas dimensiones han sido divididas en una serie de indicadores utilizados para cuantificar la variable de estudio.

La Tabla 5 muestra el proceso de operacionalización de la variable dependiente. Luego de realizar una revisión de la literatura, se consideró como dimensión de la variable dependiente el rendimiento del modelo. Esto ayudará a cuantificar la reducción del error y la precisión del modelo al evaluar la similitud entre los valores predichos y reales.

Tabla 4*Operacionalización de la variable independiente*

Variable	Dimensiones	Indicadores	Ítem	Técnica e Instrumento de recolección de datos
Modelo de ANN	Conjunto de datos	Conjunto de entrenamiento	Base de datos	Recolección y análisis de datos secundarios y Ficha de registro
		Conjunto de validación		
		Conjunto de prueba		
	Hiperparámetros asociados al modelo	Número de entradas	X	Análisis del contenido y Ficha bibliográfica
		Número de salidas	y	
		Número de capas ocultas	Num layers	Observación y Ficha de registro
		Número de neuronas ocultas	Units 1; Units 2	
		Función de activación	Activation	
	Hiperparámetros asociados al descenso del gradiente	Número de épocas	Epochs	Observación y Ficha de registro
		Tasa de aprendizaje	Learning rate	
Optimizador		Optimizer		
Tamaño de lote		Batch size		

Tabla 5*Operacionalización de la variable dependiente*

Variable	Dimensiones	Indicadores	Ítem	Técnica e Instrumento de recolección de datos	
Predicción de las propiedades mecánicas del suelo	Rendimiento del modelo	MSE	Entrenamiento	Observación y Ficha de registro	
			Validación		
			Prueba		
		MAE	Entrenamiento	Observación y Ficha de registro	
			Validación		
			Prueba		
		Coeficiente de correlación	R	Entrenamiento	Observación y Ficha de registro
				Validación	
				Prueba	

2.4. Técnicas e instrumentos de recolección de datos, validez y confiabilidad

La recolección de datos es fundamental para un estudio científico. Este paso implica la creación de una estrategia exhaustiva para los procesos que ayudarán a la recopilación de datos, teniendo en cuenta los objetivos del estudio. Esta estrategia implica conocer las fuentes, las técnicas e instrumentos de recogida de datos que se utilizarán, los procesos de recogida de datos y cómo se procesarán los datos para su análisis (Hernández-Sampieri & Mendoza Torres, 2018).

2.4.1. Técnicas de recolección de datos

La recopilación y análisis de datos secundarios se utilizó como técnica inicial de recogida de datos. Según Hernández-Sampieri y Mendoza Torres (2018), esta técnica implica la recopilación y el análisis de datos secundarios, como la revisión de documentos, archivos o registros que fueron recopilados por otros investigadores. El uso de esta técnica en este estudio consistió en recopilar y ordenar la información de los archivos técnicos que sirvieron como variables para la construcción del modelo.

El segundo método de recogida de datos fue el análisis de contenido. Según Hernández-Sampieri y Mendoza Torres (2018), esta técnica consiste en examinar cualquier tipo de información, ordenarla en categorías y subcategorías para, finalmente, analizarlas. En esta investigación, la técnica fue utilizada para hacer una revisión de diversos artículos científicos, con la finalidad de analizar las variables de entrada y salida que se suelen utilizar.

La observación fue la última estrategia de recogida de datos empleada. Esta técnica consiste en hacer un registro sistemático de diferentes escenarios observables. Posteriormente, los datos registrados son organizados a través de categorías y subcategorías (Hernández-Sampieri & Mendoza Torres, 2018; Martínez Mediano, 2014). Durante el desarrollo de las ANN y la predicción de las propiedades del suelo, se utilizó la observación para crear un registro visual de los resultados obtenidos.

2.4.2. Instrumentos de recolección de datos

Para empezar, se combinó la ficha de registro de datos junto con la técnica de recogida y análisis de datos secundarios. En la ficha se registró los resultados de las pruebas de laboratorio de suelo de todos los expedientes técnicos a los que se tuvo acceso. En segundo lugar, la ficha bibliográfica se utilizó como instrumento de la técnica del análisis del contenido y en ella se registró las variables de entrada y salida más utilizadas, según la literatura. Para la técnica de la observación se realizaron dos fichas de registro, la primera, para registrar los resultados al entrenar las ANN y, la segunda, para los resultados del rendimiento de las ANN.

Por otro lado, es importante garantizar que los datos recolectados en una investigación sean válidos y confiables. Para ello, los instrumentos de recogida de datos se consideran válidos cuando miden la variable que el investigador quiere evaluar (Hernández-Sampieri & Mendoza Torres, 2018). Así pues, los datos recolectados en este estudio para construir la base de datos son válidos, puesto que fueron obtenidos de proyectos de carreteras aprobados por el gobierno; los mismos que se realizan siguiendo cada una de las normas establecidas para los ensayos de laboratorio.

Para establecer el dominio de la variable modelo de ANN y la variable predicción de las propiedades mecánicas del suelo, se revisó 7 libros sobre el tema. También se revisó 47 artículos publicados en revistas científicas entre 2015 y 2021. Se pudo encontrar en la literatura que las dimensiones para la variable independiente fueron: (a) conjunto de datos, (b) hiperparámetros asociados al modelo y (c) hiperparámetros asociados al descenso del gradiente; y para la variable dependiente: rendimiento del modelo (Hernández-Sampieri & Mendoza Torres, 2018).

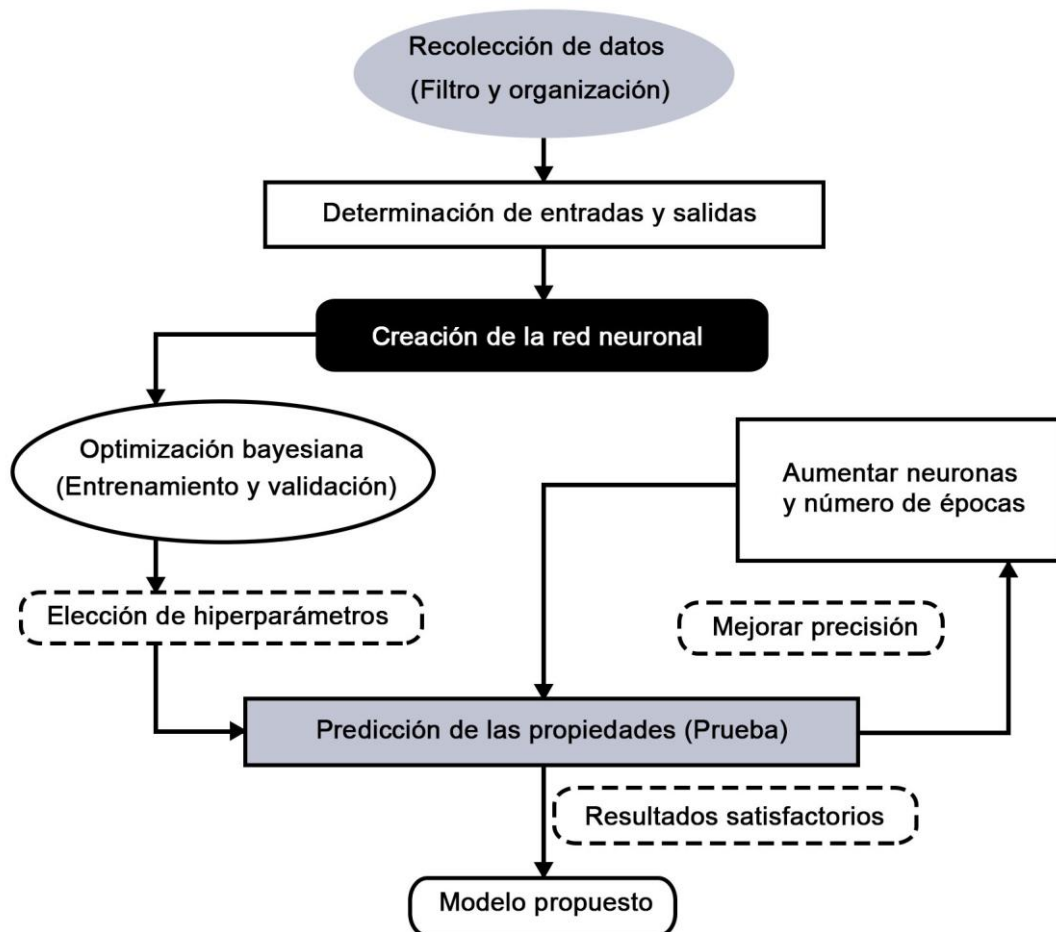
Un instrumento es confiable cuando da los mismos resultados luego de haber sido aplicado varias veces a la misma persona u objeto (Hernández-Sampieri & Mendoza Torres, 2018). Teniendo en cuenta estos conceptos, los datos para la base de datos son confiables ya que son de proyectos de carreteras ya ejecutados. Asimismo, los datos recolectados para el resto de dimensiones también son confiables. Puesto que, para el desarrollo de los modelos se aplicó 3 veces el mismo procedimiento y así evaluar la variabilidad de los datos.

2.4.3. Procedimientos para la recolección de datos

En la Figura 16 se observa el proceso de la investigación. Como primer paso, se realizó la búsqueda de expedientes técnicos. Se ingresó al buscador público del SEACE y se filtró la búsqueda a obras de la región Lambayeque publicadas entre los años 2015 y 2021; aunque, solo se encontró los expedientes técnicos de las obras publicadas del año 2019 hacia adelante. Se identificó a aquellos proyectos de pavimentación y se verificó que su expediente técnico se encontrara disponible.

Figura 16

Procedimiento seguido para el desarrollo de las ANN



Luego de descargar los expedientes técnicos, se aplicó la primera ficha de registro de datos (ver Anexo 1). En la ficha se registró la información para el ítem base de datos, como se indica en la operacionalización de variables. Los datos

fueron ordenados, revisados y se codificaron siguiendo la Tabla 6. Se obtuvieron 12 variables de los ensayos y 2 variables de la clasificación del suelo. Estas fueron analizadas para ser usadas como entradas y salidas de los modelos. Esta primera ficha de registro sirvió para crear la base de datos que se utilizaría para entrenar los modelos.

Tabla 6

Variables consideradas para los modelos de ANN

Ensayo	Variables	Símbolo
Contenido de humedad	Porcentaje de humedad	H
Contenido de sales	Porcentaje de sales	S
Análisis granulométrico	Porcentaje de grava	G
	Porcentaje de arena	A
	Porcentaje de finos	F
Límites de Atterberg	Límite líquido	LL
	Límite plástico	PL
	Índice de plasticidad	PI
Clasificación	USCS	SUCS
	AASHTO	AASHTO
Compactación	Máxima densidad seca	MDD
	Óptimo contenido de humedad	OMC
CBR	Al 95 % de la máxima densidad seca	CBR95
	Al 100 % de la máxima densidad seca	CBR100

Nota. G = suelo retenido en el tamiz n.º 4 tras pasar el tamiz de 3 in.; A = suelo retenido en el tamiz n.º 200 tras pasar el tamiz n.º 4; F = suelo que pasa el tamiz n.º 200.

Como segundo paso, se revisó diversos artículos científicos con el fin de analizar qué variables de entrada y salida utilizaron en sus ANN. En total se registró la información de 10 artículos y, para ello, se empleó la ficha bibliográfica (ver Anexo 2). La ficha estuvo organizada por artículos consultados y por propiedades del suelo ya mencionadas en el paso anterior. En ella se recolectó los datos para el ítem “X” (número de entradas) y el ítem “y” (número de salidas). Si un estudio utilizaba una determinada propiedad se marcaba con el símbolo dependiendo de su uso como variable de entrada o salida.

Teniendo los datos y sabiendo qué entradas y salidas utilizar, se importó la base de datos desde Spyder y se definió las entradas y salidas del modelo. Se realizó el preprocesamiento del conjunto de datos, para ello se utilizó el estimador MinMaxScaler. Hay que mencionar que este procedimiento se hizo para cada una de las propiedades del suelo utilizadas como variables. El estimador se importó desde el programa Spyder utilizando el código que se muestra a continuación:

```
from sklearn.preprocessing import MinMaxScaler
```

El MinMaxScaler escala cada dato utilizando la ecuación (2.1). Donde x es el valor que se desea escalar, x_{min} el valor mínimo de todo el conjunto de datos y x_{max} el valor máximo de todo el conjunto de datos.

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2.1)$$

En el tercer paso, se realizó una ficha de registro para encontrar la mejor arquitectura (ver Anexo 3). Los datos se dividieron en dos partes: un conjunto de entrenamiento y un conjunto de prueba. Se realizó el código para entrenar los modelos y se definió la cuadrícula de búsqueda de hiperparámetros. Se realizaron 100 iteraciones por cada propiedad del suelo a predecir utilizando la optimización bayesiana y la validación cruzada. Cada iteración se entrenó 3 veces ($K = 3$) con la misma configuración. Al finalizar, se exportó los resultados de Spyder y se aplicó la ficha para ordenar y registrar los datos sobre los ítems: num layers, units 1 y units 2, activation, epochs, learning rate, optimizer y batch size. Es decir, se registró los hiperparámetros usados por cada modelo y su error cometido.

Siguiendo con el cuarto paso, se seleccionó los mejores modelos para cada propiedad y se hicieron predicciones de los parámetros de compactación y CBR. Para ello, se creó una ficha de registro de datos para comprobar el rendimiento de los modelos (ver Anexo 4). En ella se recogieron datos como el MSE, el MAE y el R. La recolección de los datos en esta etapa fue importante puesto que permitió comparar las predicciones del modelo con los datos reales.

El registro de los datos se realizó a lo largo del estudio con el programa Excel. Los datos registrados en la primera ficha fueron corroborados con los expedientes técnicos. Este procedimiento se realizó para asegurar que los datos fuesen los correctos. Lo mismo se buscó con los datos recolectados en las fichas número tres y cuatro. Estos datos fueron exportados tal y como se obtuvieron de Spyder; se guardaron en formato de archivo de valores separados con comas. Todo esto ayudó a separar los resultados que el recojo de información errónea podría imponer a la teoría frente a la teoría real que emerge de los datos.

2.5. Procedimiento de análisis de datos

Toda la información recopilada se organizó en tablas, empezando por la recogida de muestras para el conjunto de datos y terminando con los resultados. Asimismo, se utilizó Excel y R Studio para analizar los datos recolectados. El uso de estos programas ayudó a explorar, analizar y visualizar los datos recolectados en la etapa anterior.

Para obtener los resultados de este estudio fue esencial utilizar diversas técnicas de análisis estadístico para analizar los datos. El procedimiento y el momento en que fueron aplicadas cada una de las técnicas se describe más adelante. En resumen, las técnicas usadas fueron: (a) media, (b) rango, (c) desviación estándar, (d) MSE, (e) MAE, (f) R y (g) análisis de dispersión.

2.5.1. Entradas y salidas del modelo

Como ya se vio anteriormente, el estudio de la literatura se utilizó para investigar la tendencia del uso de las propiedades como variables de los modelos de ANN. A modo de ayuda para identificar que variables de entrada y salida usar, se realizó

un análisis de dispersión entre las propiedades. Es decir, se identificó si existía una correlación entre H, S, G, A, F, LL, PL y PI con los parámetros de compactación y CBR. Esto con la finalidad de ver si alguna propiedad influye en la predicción.

2.5.2. División de los datos

La media se utilizó como medida de tendencia central para comparar los conjuntos de entrenamiento y prueba, con el fin de colocar los datos en una misma escala de medición para cada propiedad analizada. La media aritmética, que se calcula dividiendo el total de todos los valores entre la cantidad de observaciones, es una de las estadísticas más utilizadas. La expresión (2.2) se utiliza para calcular la media; donde x_i son los valores del conjunto de datos y n es la cantidad total de datos (Hernández-Sampieri & Mendoza Torres, 2018).

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} \quad (2.2)$$

También se analizó la dispersión entre los conjuntos utilizados para entrenar y probar el modelo. El rango de valores se calculó como la diferencia entre el valor más alto y el mínimo como medida de variabilidad, como se muestra en la ecuación (2.3). La desviación estándar se utilizó para analizar la diferencia media entre los conjuntos utilizados para entrenar y probar el modelo, para ello se utilizó la ecuación (2.4). Este análisis permitió explorar los datos y ver la variación existente luego de dividirlos. Ahora bien, es importante mencionar que este análisis se realizó a los datos sin escalar.

$$Rango = (Max) - (Min) \quad (2.3)$$

$$\sigma = \sqrt{\frac{\sum_i^n (x_i - \bar{X})^2}{n}} \quad (2.4)$$

2.5.3. Determinación de la arquitectura del modelo

Durante el entrenamiento, ajuste de hiperparámetros y al momento de evaluar el buen comportamiento de los modelos se utilizó la expresión (2.5) que corresponde al error cuadrático medio. También se utilizó como métrica el error absoluto medio, como se ve en la ecuación (2.6). En esta fórmula, y_i es el valor predicho, x_i es el valor real y n el número de datos.

$$MSE = \frac{\sum_{i=1}^n (y_i - x_i)^2}{n} \quad (2.5)$$

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} \quad (2.6)$$

2.5.4. Prueba del modelo

Para probar los modelos de ANN se utilizó la correlación de Pearson, la cual utiliza la ecuación (2.7). Esta comparación se realizó entre los valores previstos y los reales por los modelos. También ayudó a asegurar que los modelos desarrollados en este estudio para estimar la compactación del suelo y los valores de CBR eran válidos.

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (2.7)$$

2.6. Criterios éticos

En este estudio se trabajó con el reporte de Belmont. Donde según Belmont (1979), son tres los principios éticos que rigen una investigación, los mismos que se detallan a continuación:

- **información:** se brinda la información necesaria sobre el procedimiento y los fines de esta investigación.
- **comprensión:** se asegura que la información ha sido comprendida.

2.7. Criterios de rigor científico

Este estudio se realizó teniendo en cuenta los siguientes criterios de rigor científico:

- **credibilidad:** los datos recogidos en la investigación tienen concordancia con la realidad, es decir, los resultados obtenidos presentan una relación con lo que se observó.
- **neutralidad:** no se manipularon los resultados obtenidos en esta investigación.
- **relevancia:** permite evaluar si se han cumplido los objetivos del estudio.

III. RESULTADOS

En la sección 3.1 se presentan los resultados para ver en qué medida las ANN son eficaces para predecir el CBR y los parámetros de compactación del suelo. La evaluación del desempeño de los modelos se midió en términos de disminución del error de predicción. Asimismo, se presenta los resultados para ajustar los hiperparámetros de los modelos de ANN, tal como el número de neuronas, capas ocultas, entre otros. Se escogió a los mejores modelos de ANN y se hizo predicciones utilizando datos desconocidos para el modelo. Luego, se comparan los valores predichos con los reales para evaluar el rendimiento. Finalmente, en la sección 3.2 se hace una discusión de los resultados.

3.1. Resultados en tablas y figuras

Los resultados permiten demostrar la hipótesis planteada en esta investigación. En ella se sostiene que los modelos de ANN desarrollados predicen las propiedades mecánicas del suelo de la región Lambayeque con una precisión del 80 %. Es decir, los valores predichos de la MDD, OMC, CBR al 95 % de MDD y CBR al 100 % de MDD se correlacionan con sus valores reales.

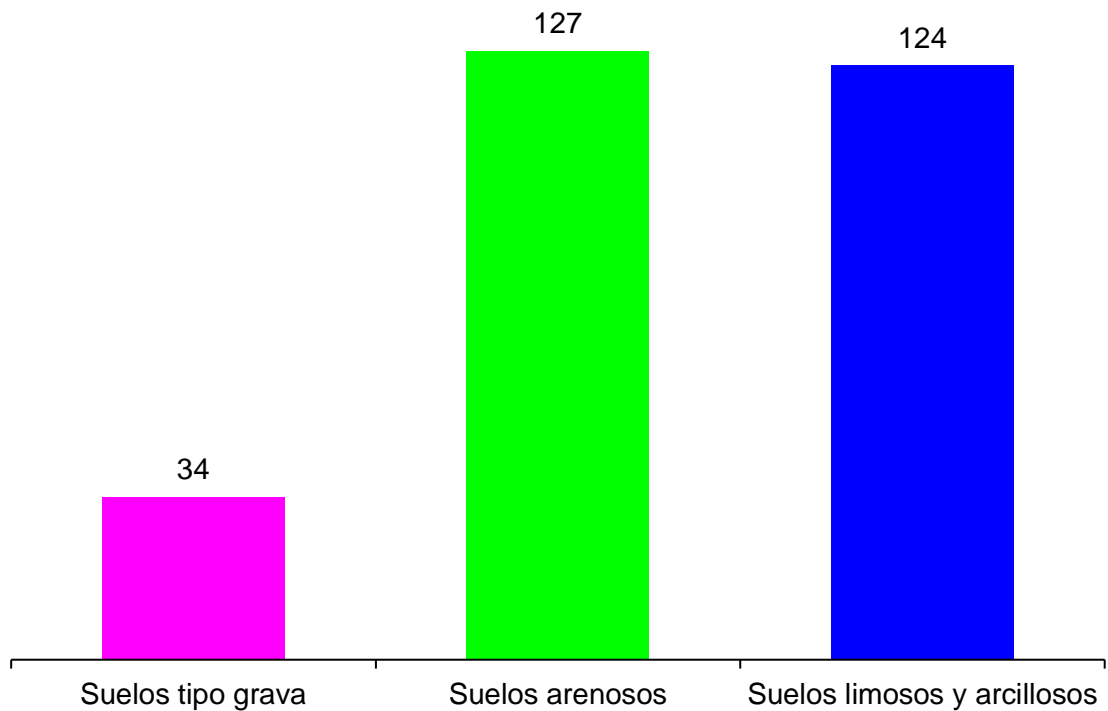
Para describir los resultados de este estudio se utilizan tablas, gráficos de líneas y diagramas de barras. Los resultados se presentan en tres secciones en este apartado. En la primera parte se presenta la exploración del conjunto de datos. En la segunda, se presenta la selección de entradas y salidas para los modelos. Luego la división del conjunto de datos. Luego se presentan los modelos para predecir las propiedades; esta parte abarca el entrenamiento, validación y etapa de prueba, utilizando los mejores modelos para realizar predicciones. Posteriormente, se busca mejorar la precisión de los modelos ya desarrollados. Finalmente, se muestra las correlaciones entre las propiedades predichas por los modelos propuestos y las propiedades reales.

3.1.1. Exploración del conjunto de datos

Se utilizaron 285 muestras de suelo para el desarrollo de los modelos de ANN. El Anexo 5 indica las cualidades de las muestras de suelo utilizadas como conjunto de datos. La Figura 17 muestra cómo estuvo distribuido el conjunto de datos según el tipo de suelo. Del total de ejemplos del conjunto de datos, según su clasificación USCS, 34 (11.93 %) fueron de suelos de tipo grava, 127 (44.56 %) fueron suelos arenosos y 124 (43.51 %) fueron muestras de suelo limosos y arcillosos.

Figura 17

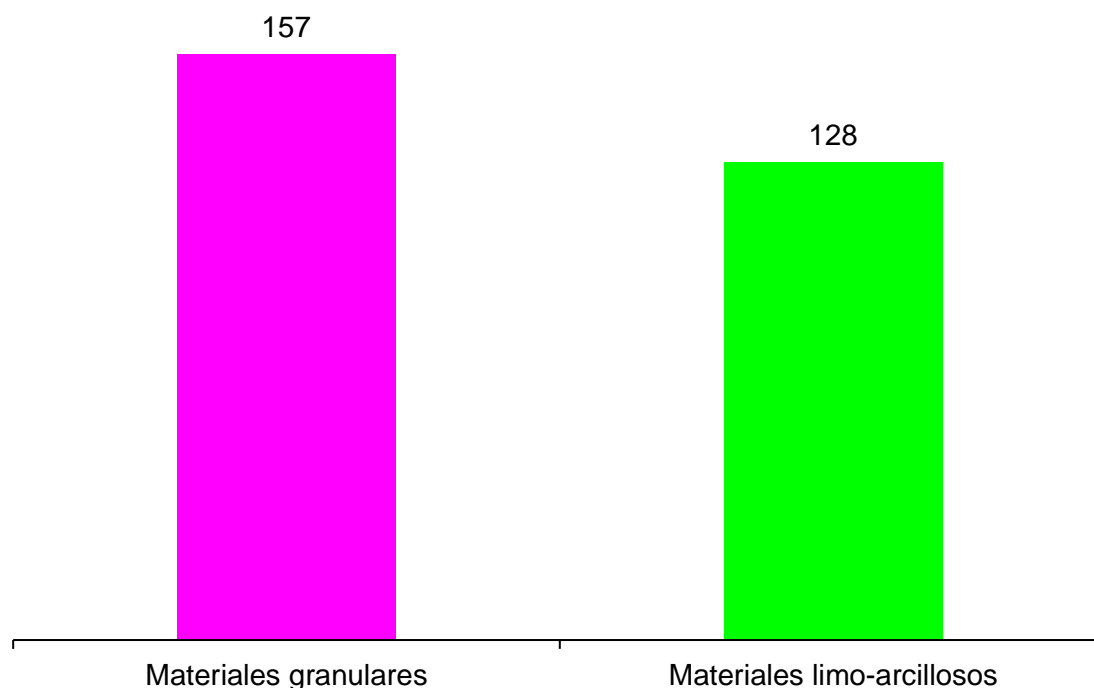
Distribución del conjunto de datos según su clasificación USCS



La Figura 18 muestra la distribución del número de ejemplos del conjunto de datos por tipo de suelo. Según el sistema AASTHO, 157 (55.09 %) fueron materiales granulares y 128 (44.91 %) fueron materiales limo-arcillosos. Se puede observar que el número de muestras clasificados como suelos finos es similar al número de ejemplos según el sistema de clasificación USCS.

Figura 18

Distribución del conjunto de datos según su clasificación AASTHO



De todo el conjunto de datos, el contenido de humedad de las muestras estuvo entre 1.10 % y 36.07 %. El contenido de sales estuvo entre 0 % y 2.52 %. El contenido de grava de las muestras estuvo entre 0 % y 81.9 %. El contenido de arena de 1.97 % y 100 %. Tuvieron un contenido de finos entre 0% y 97.69 %. LL de las muestras estuvo entre 0 % y 59.58 %. PL entre 0 % y 27.93 %. El PI estuvo entre 0 % y 37.44 %. MDD de las muestras de suelo estuvo entre 1.558 g/cm³ y 2.21 g/cm³. OMC entre 5.58 % y 22.33 %. CBR al 95 % de la MDD entre 2.81 % y 72.00 %. CBR al 100 % de la MDD entre 4.20 % y 86.00 %.

Se descargó un total de 9 expedientes técnicos publicados en el SEACE en el año 2019, 26 expedientes técnicos publicados en el año 2020 y 1 expediente técnico del 2021. De la provincia de Chiclayo, las muestras de suelo pertenecieron a los distritos de Chiclayo, Chongoyape, José Leonardo Ortiz, Pimentel, Pátapo, Pomalca, Puerto Eten, Reque, Santa Rosa, Tumán y Zaña. En la provincia de Ferreñafe, las muestras fueron de los distritos de Ferreñafe y Pítipo. Para la provincia de Lambayeque, la muestra de suelo fue del distrito de Lambayeque,

Mochumí, Mórrope, Motupe, San José y Túcume. La Tabla 7 muestra de manera detallada el número de expedientes técnicos descargados, así como, el número de muestras y ensayos recopilados de los años ya mencionados.

Tabla 7

Información recopilada del SEACE de los años 2019, 2020 y 2021

Lugar	Expedientes	Muestras	Ensayos
2019			
Chiclayo	6	92	552
Ferreñafe	2	19	114
Lambayeque	1	6	36
Total	9	117	702
2020			
Chiclayo	15	87	522
Ferreñafe	2	16	96
Lambayeque	9	44	264
Total	26	147	882
2021			
Chiclayo	0	0	0
Ferreñafe	0	0	0
Lambayeque	1	21	126
Total	1	21	126

Nota. Ensayos por muestra: a) humedad, b) sales, c) granulometría, d) límites de consistencia, e) compactación, y f) CBR.

3.1.2. Entradas y salidas del modelo

Para los modelos se consideró una variable de salida debido a que se quiere resolver un problema de regresión mediante modelos de ANN. La Tabla 8 muestra los modelos finales y las variables de entrada que se utilizaron. En general, se consideraron 8 variables de entrada en los modelos MARK 1, MARK 2, MARK 5 y MARK 6 (ver Figura 19). Por el contrario, se usó 10 variables de entrada en los modelos MARK 3 y MARK 4 (ver Figura 20). Esta variación se hizo para analizar el efecto del número de variables en el resultado.

Tabla 8

Entradas y Salidas utilizadas

Modelo	H	S	G	A	F	LL	PL	PI	MDD	OMC	CBR95	CBR100
MARK 1	X	X	X	X	X	X	X	X	-	-	-	y
MARK 2	X	X	X	X	X	X	X	X	-	-	y	-
MARK 3	X	X	X	X	X	X	X	X	X	X	-	y
MARK 4	X	X	X	X	X	X	X	X	X	X	y	-
MARK 5	X	X	X	X	X	X	X	X	y	-	-	-
MARK 6	X	X	X	X	X	X	X	X	-	y	-	-

Nota. X = propiedades usadas como entradas; y = propiedades utilizadas como salidas.

La selección de las variables se realizó según la información de los datos. Se utilizó un análisis de correlación para encontrar las propiedades que se relacionaban con el valor a predecir y, por ende, no debían faltar en el modelo. En segundo lugar, se realizó una revisión de la literatura para determinar la tendencia actual del uso de las propiedades como variables de entrada (ver Tabla 9). Estas incluyen variables como H, S, G, A, F, LL, PL, PI, MDD, OMC, CBR y otras propiedades.

Figura 19

Arquitectura inicial de los modelos de ANN con 8 variable de entrada

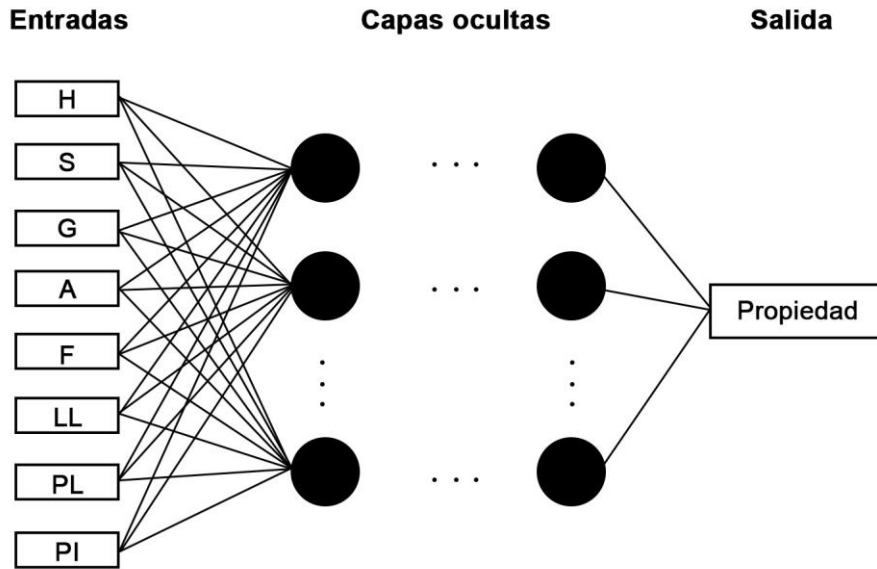


Figura 20

Arquitectura inicial de los modelos de ANN con 10 variables de entrada

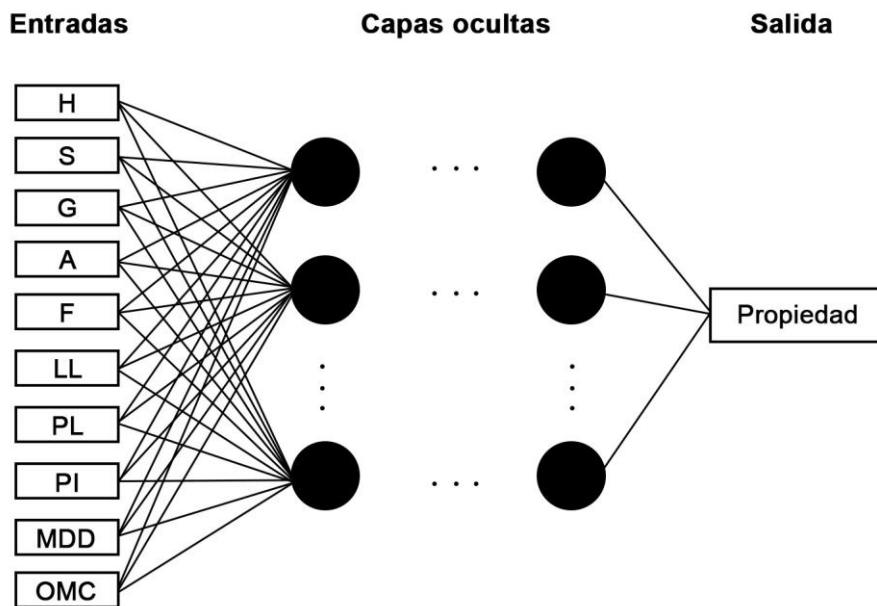


Tabla 9*Entradas y salidas de modelos según literatura*

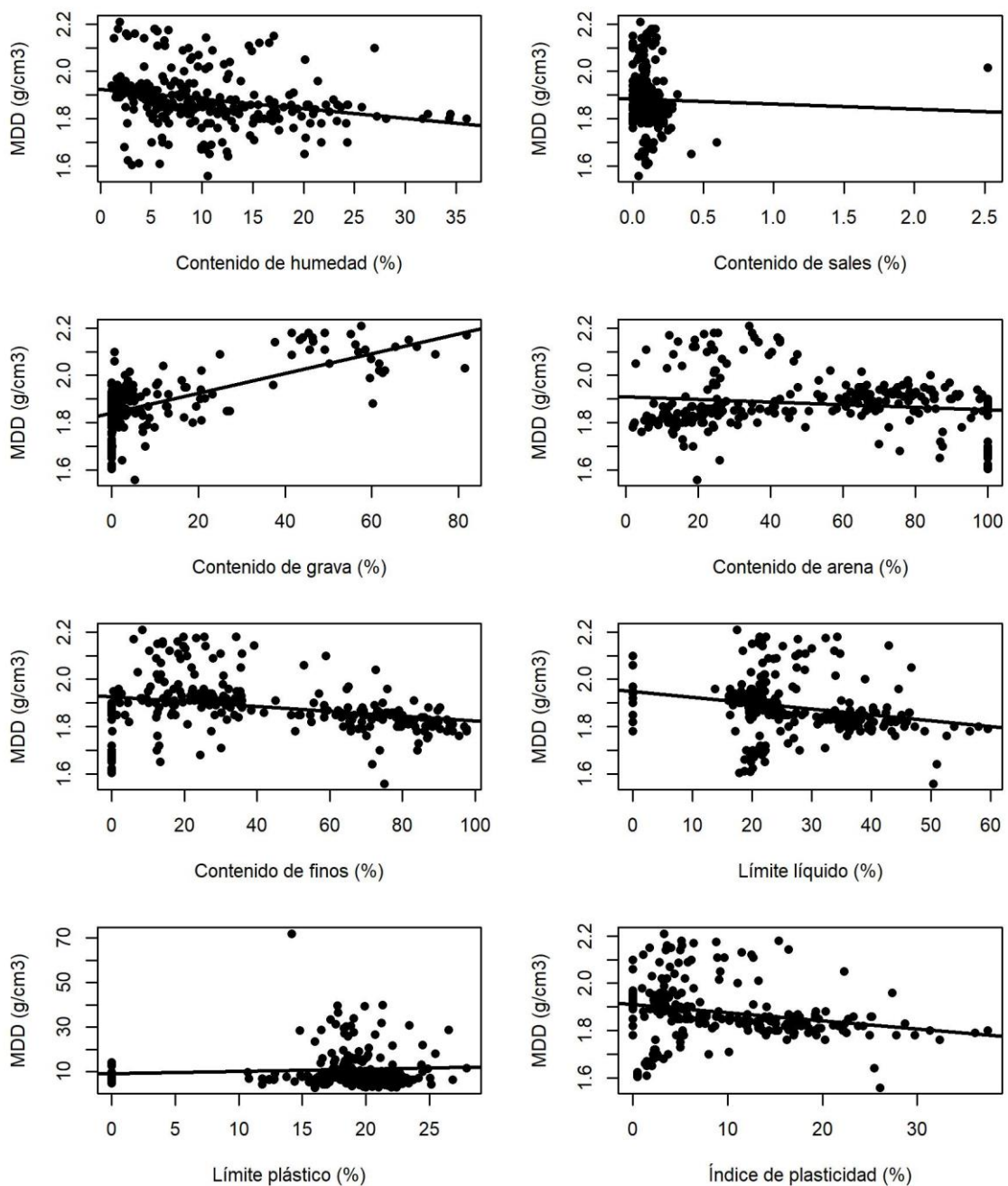
Estudio	H	S	G	A	F	LL	PL	PI	O	MDD	OMC	CBR
Tenpe y Patel (2020)	X	-	X	X	X	X	X	X	-	X	X	y
Alam et al. (2020)	-	-	-	-	-	X	X	X	X	X	X	y
Rajakumar y Reddy babu (2020)	-	-	-	-	-	X	X	-	X	X	X	y
Fikret Kurnaz y Kaya (2019)	-	-	X	X	X	-	-	-	-	X	X	y
Ghorbani y Hasanzadehshooiili (2018)	-	-	-	-	X	-	-	-	X	-	-	y
Suthar y Aggarwal (2018)	-	-	-	-	-	-	-	-	X	X	X	y
Erzin y Turkoz (2016)	X	-	-	-	-	-	-	-	X	X	-	y
Albusoda <i>et al.</i> (2020)	X	X	-	X	X	X	X	X	X	y	-	-
Ardakani y Kordnaeij (2019)	-	-	-	X	X	X	X	-	-	y	y	y
Manrique Dávila (2019)	-	-	X	X	X	X	X	X	X	y	y	y

Nota. O = otras propiedades; X = propiedades usadas como entradas; y = propiedades utilizadas como salidas.

La Figura 21 muestra los diagramas de dispersión entre las variables de salida MDD y las posibles variables de entrada a utilizar en los modelos de ANN para esta propiedad. Según el análisis de correlación, se encontró que la variable de salida MDD tuvo una correlación significativa con las variables H, S, G, F, LL, PL y PI. Asimismo, no hubo una correlación significativa con A (ver Anexo 6).

Figura 21

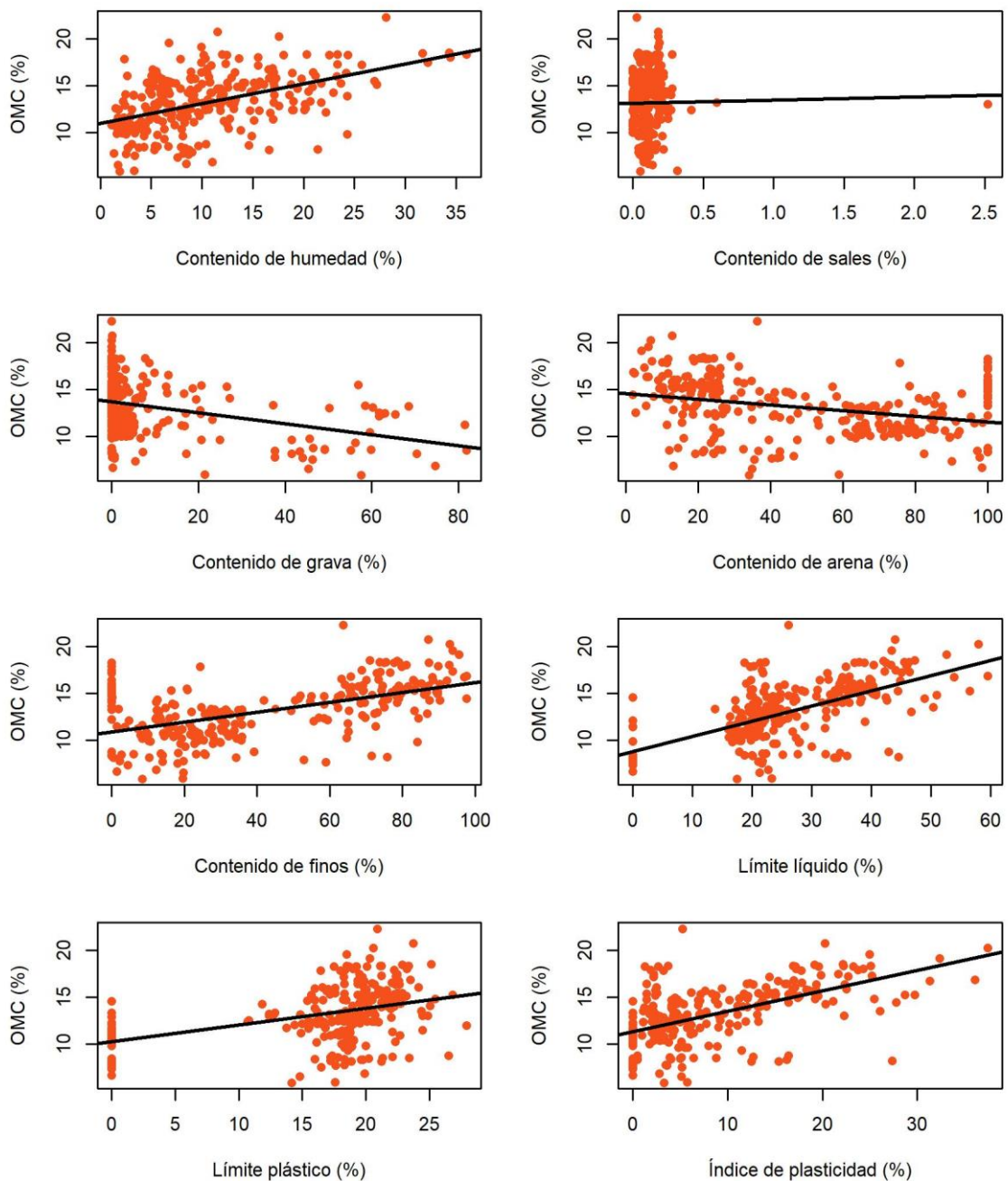
Diagrama de dispersión para MDD y las variables de entrada



La Figura 22 muestra los diagramas de dispersión de la variable de salida OMC y sus posibles variables de entrada a utilizar en los modelos de ANN. Según el análisis de correlación, se encontró una relación significativa de la variable de salida con H, G, A, F, LL, PL y PI. Asimismo, no se encontró una relación significativa con S (ver Anexo 6).

Figura 22

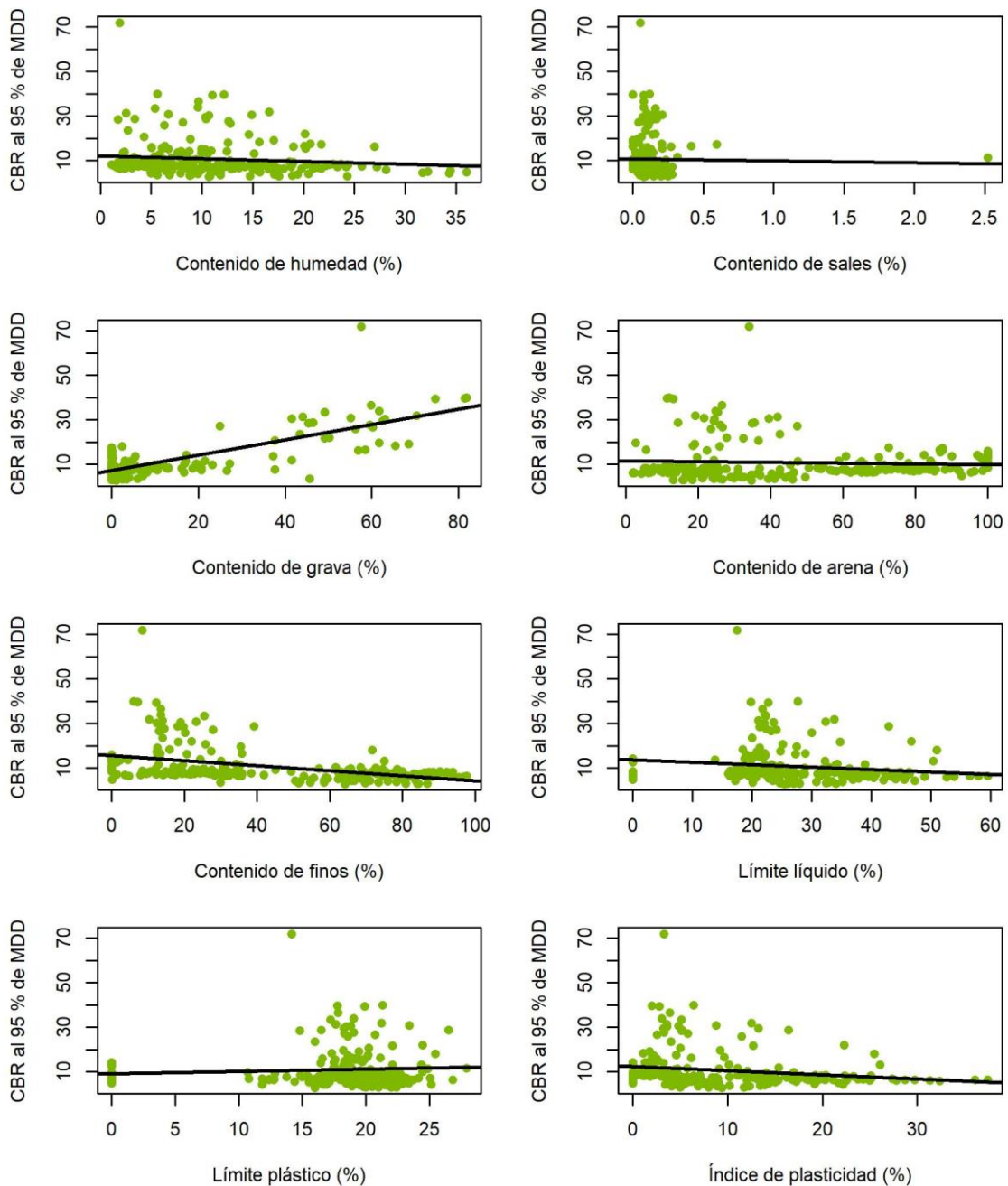
Diagrama de dispersión para OMC y las variables de entrada



La Figura 23 muestra la dispersión entre el CBR al 95% de MDD y las posibles variables de entrada a utilizar en los modelos de ANN para esta propiedad. En el análisis de correlación se encontró una correlación significativa de la variable de salida con H, S, G, A, F, LL y PI. Asimismo, no se encontró una relación significativa con PL (ver Anexo 6).

Figura 23

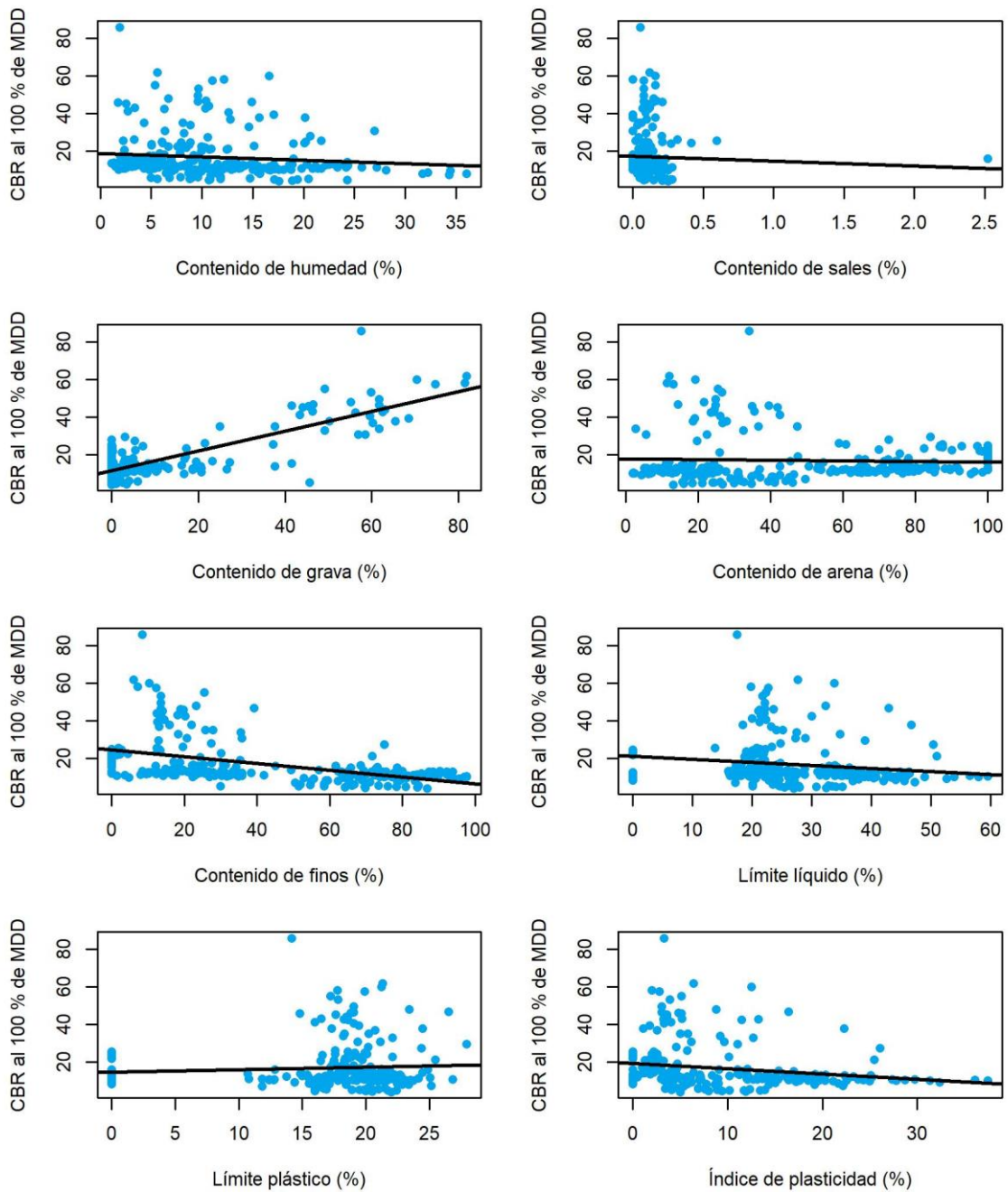
Diagrama de dispersión para el CBR al 95 % de la MDD y las variables de entrada



En la Figura 24 se ve el gráfico de dispersión entre la variable de salida CBR al 100 % y las posibles variables de entrada del modelo de ANN para predecir esta propiedad. Según el análisis de correlación, se encontró una correlación significativa de esta variable de salida con H, S, G, A, F, LL y PI. Asimismo, no se encontró relación con el PL (ver Anexo 6).

Figura 24

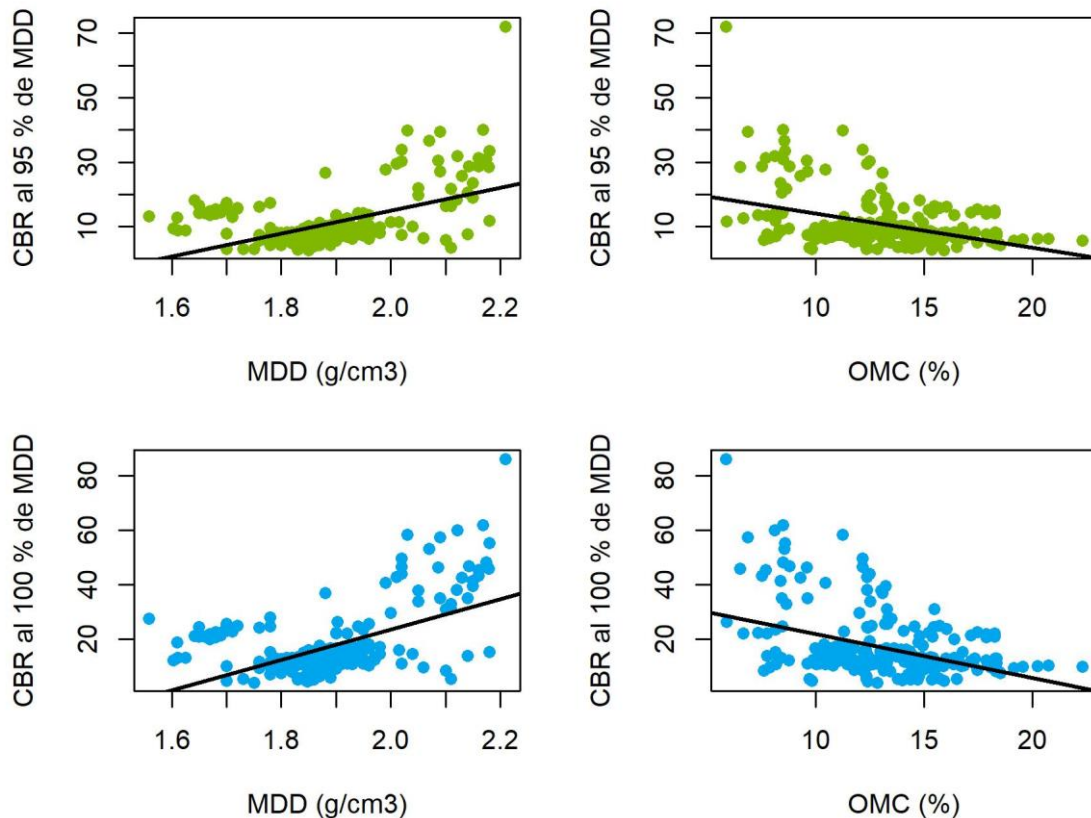
Diagrama de dispersión para el CBR al 100 % de MDD y las variables de entrada



En la Figura 25 se muestra la dispersión entre el CBR con la MDD y OMC. Según los resultados del análisis de correlación, existe una correlación significativa del CBR (al 95 % de MDD y al 100 % de MDD) con la MDD y el OMC (ver Anexo 6).

Figura 25

Diagrama de dispersión para el CBR y la compactación



3.1.3. División del conjunto de datos

Los datos fueron divididos en un conjunto de entrenamiento y un conjunto de prueba que, junto a la validación cruzada, permitieron ajustar los modelos. La división del conjunto de datos se hizo de manera aleatoria. La Tabla 10 y Tabla 11 muestran el resultado de las pruebas estadísticas que se hizo a ambos conjuntos. Se realizó un análisis de medida central y medias de variabilidad. Estos resultados ayudaron a identificar si existía una representatividad de los ejemplos en ambos conjuntos, la cual se verá más adelante.

Tabla 10*Estadísticas del conjunto de entrenamiento y prueba para las variables de entrada*

VARIABLES	Media	σ	Mínimo	Máximo	Rango
Contenido de humedad, H (%)					
Entrenamiento	9.90	6.70	1.10	36.07	34.97
Prueba	11.56	8.07	1.30	32.24	30.94
Contenido de Sales, S (%)					
Entrenamiento	0.11	0.18	0.00	2.52	2.52
Prueba	0.11	0.09	0.00	0.60	0.60
Contenido de Grava, G (%)					
Entrenamiento	10.54	19.46	0.00	81.90	81.90
Prueba	7.64	16.16	0.00	62.45	62.45
Contenido de Arena, A (%)					
Entrenamiento	47.17	30.86	1.97	100.00	98.03
Prueba	48.73	31.20	7.19	100.00	92.81
Contenido de Finos, F (%)					
Entrenamiento	42.30	30.37	0.00	97.69	97.69
Prueba	43.62	32.02	0.00	92.81	92.81
Límite líquido, LL (%)					
Entrenamiento	26.75	10.89	0.00	59.58	59.58
Prueba	26.68	11.30	0.00	56.55	56.55
Límite plástico, PL (%)					
Entrenamiento	16.17	7.82	0.00	27.93	27.93
Prueba	16.27	7.60	0.00	26.84	26.84
Índice de plasticidad, PI (%)					
Entrenamiento	8.25	8.06	0.00	37.44	37.44
Prueba	8.47	7.87	0.00	29.71	29.71

Tabla 11*Estadísticas del conjunto de entrenamiento y prueba para las variables de salida*

Variabes	Media	σ	Mínimo	Máximo	Rango
Máxima densidad seca, <i>MDD</i> (g/cm ³)					
Entrenamiento	1.89	0.12	1.56	2.21	0.65
Prueba	1.87	0.12	1.60	2.18	0.58
Óptimo contenido de humedad, <i>OMC</i> (%)					
Entrenamiento	13.09	2.99	5.85	22.33	16.48
Prueba	13.33	3.09	6.50	18.47	11.97
CBR al 95% de la MDD, <i>CBR95</i> (%)					
Entrenamiento	10.98	8.58	2.81	72.00	69.19
Prueba	9.80	6.13	3.16	29.50	26.34
CBR al 100% de la MDD, <i>CBR100</i> (%)					
Entrenamiento	17.41	12.51	4.20	86.00	81.80
Prueba	15.30	9.22	4.60	45.90	41.30

Para asegurar que los conjuntos de entrenamiento y prueba sean representativos, en esta investigación se decidió realizar una comparación de las medias de ambos conjuntos. En los resultados de las tablas anteriores se observa que las medias de los conjuntos de entrenamiento y prueba son similares. La finalidad es asegurar, por ejemplo, que después de dividir los datos no solo hubiera muestras clasificadas como grava en el conjunto de entrenamiento y muestras clasificadas como arenas, limos y arcillas en el conjunto de prueba. Sino que existan ejemplos de los tres tipos de muestras en ambos conjuntos.

En la Tabla 12 muestra los resultados de la prueba U de Mann-Whitney, que reveló que no hubo una diferencia significativa entre los conjuntos de entrenamiento y prueba para H, S, G, A, F, LL, PL y PI. Con estos resultados se concluye que el conjunto de entrenamiento y el conjunto de prueba para las variables de entrada son representativos.

Tabla 12

Prueba U de Mann-Whitney para la igualdad de medias entre los conjuntos de entrenamiento y prueba para las variables de entrada

Variables de entrada	W	Sig. (bilateral)	99 % de intervalo de confianza de la diferencia	
			Inferior	Superior
H	5850.0	0.245	-4.070	1.420
S	6391.5	0.849	-0.027	0.020
G	7427.0	0.090	0.000	1.190
A	6256.0	0.664	-11.520	8.300
F	6296.5	0.718	-11.570	9.030
LL	6506.5	0.989	-2.920	2.810
PL	6494.0	0.995	-1.320	1.400
PI	6320.0	0.749	-2.540	2.080

Como se observa en la Tabla 13, no hubo diferencias significativas entre los conjuntos de entrenamiento y de prueba para MDD, según la prueba U de Mann-Whitney. Asimismo, tampoco hubo una diferencia entre los conjuntos de entrenamiento y conjunto de prueba para el CBR al 95 % de MDD y CBR al 100 % de MDD. Por lo que ambos conjuntos son representativos.

Tabla 13

Prueba U de Mann-Whitney para la igualdad de medias entre los conjuntos de entrenamiento y prueba para las variables de salida

Variables de salida	W	Sig. (bilateral)	99 % de intervalo de confianza de la diferencia	
			Inferior	Superior
MDD	7011	0.357	-0.020	0.050
OMC	6086	0.460	-1.630	0.870
CBR al 95 % de MDD	6889	0.483	-0.700	1.300
CBR al 100 % de MDD	6904	0.466	-0.960	2.180

3.1.4. Desarrollo de modelos de ANN para predecir el CBR al 100 % de MDD

Al finalizar se proponen un modelo para predecir el CBR al 100 % de MDD. El procedimiento consistió en hacer 100 iteraciones utilizando ocho variables de entrada y 100 iteraciones utilizando diez variables de entrada, realizar predicciones con el mejor modelo y, finalmente, evaluar el rendimiento del modelo comparando los valores reales con los predichos. De los primeros 100 se seleccionó el que tuvo mejor rendimiento y se le llamó MARK 1. De los otros 100 modelos se seleccionó el mejor modelo y se llamó MARK 3. Las características de los modelos se verán más adelante.

3.1.4.1. Ajuste de hiperparámetros

La Tabla 14 muestra los 10 mejores modelos de ANN para predecir el CBR al 100% de MDD de un total de 100 (ver Anexo 7). Los hiperparámetros se seleccionaron utilizando la optimización Bayesiana y la validación cruzada con K = 10 por 3 repeticiones. Las variables de entrada fueron H, S, G, A, F, LL, PL y PI. El modelo n.º 2 fue el mejor y tuvo un MSE de 5.47E-03. En esta investigación se hace

referencia a este modelo como MARK 1 y la lista de hiperparámetros se muestra más adelante.

Tabla 14

Mejores 10 ANN para predecir el CBR al 100% de la MDD (8 entradas)

Ranking	Modelo	Capas ocultas	Neuronas		Learning rate	MSE
			(a)	(b)		
1	2	2	5	18	1E-04	5.47E-03
2	83	2	20	20	4E-05	5.56E-03
3	65	5	18	8	4E-05	5.76E-03
4	79	2	20	20	1E-04	5.77E-03
5	35	2	20	17	2E-04	5.81E-03
6	92	6	20	20	1E-04	5.93E-03
7	31	2	20	15	3E-04	6.00E-03
8	26	2	20	5	5E-04	6.17E-03
9	36	2	20	14	1E-03	6.27E-03
10	81	6	20	20	1E-04	6.47E-03

Nota. (a) = n.º de neuronas en la primera capa oculta; (b) = n.º de neuronas en las capas ocultas restantes.

La Tabla 15 muestra los 10 mejores modelos de ANN de un total de 100 iteraciones para predecir el CBR al 100% de MDD (ver Anexo 8). Los hiperparámetros se seleccionaron utilizando la optimización Bayesiana y la validación cruzada con K = 10 por 3 repeticiones. Las variables de entrada fueron H, S, G, A, F, LL, PL, PI, MDD y OMC. El modelo n.º 82 fue el mejor modelo de ANN para predecir esta propiedad y tuvo un MSE de 4.60E-03. En esta investigación se hace referencia a este modelo como MARK 3 y su configuración se indica más adelante.

Tabla 15*Mejores 10 modelos para predecir el CBR al 100% de la MDD (10 entradas)*

Ranking	Modelo	Capas ocultas	Neuronas		Learning rate	MSE
			(a)	(b)		
1	82	1	16	—	1E-03	4.60E-03
2	30	5	11	11	2E-03	4.66E-03
3	48	1	20	—	1E-02	4.85E-03
4	17	5	18	20	3E-03	4.88E-03
5	59	5	16	20	1E-03	4.97E-03
6	71	4	10	20	2E-03	5.06E-03
7	18	6	11	20	3E-02	5.35E-03
8	51	1	15	—	3E-03	5.49E-03
9	20	6	12	18	6E-04	5.70E-03
10	11	6	16	14	6E-03	5.70E-03

Nota. (a) = n.^o de neuronas en la primera capa oculta; (b) = n.^o de neuronas en las capas ocultas restantes.

3.1.4.2. Predicciones

A continuación, se presentan los resultados obtenidos de las predicciones de los valores de CBR al 100% de MDD utilizando las ANN MARK 1 y MARK 3. Los ejemplos utilizados en el conjunto de prueba son los mismos para ambos modelos. Se hicieron 2 modelos con la finalidad de observar la variación de la precisión al aumentar o disminuir la cantidad de variables de entrenada.

MARK 1

En primer lugar, la Tabla 16 muestra la lista de hiperparámetros que utilizó el modelo de ANN MARK 1 para predecir el CBR al 100 % de MDD. Los resultados

de la optimización bayesiana produjeron esa configuración durante la búsqueda de hiperparámetros. Esto indica que utilizando estos hiperparámetros, el modelo de ANN MARK 1 produjo los mejores resultados de las 100 iteraciones realizadas empleando las ocho entradas ya mencionadas. Cabe destacar que los resultados de MSE y MAE de la tabla son los resultados obtenidos en las predicciones, es decir, durante la etapa de prueba. Estos errores cometidos por el MARK 1 están en las mismas unidades de la propiedad a predecir.

Tabla 16

Hiperparámetros del modelo de ANN MARK 1

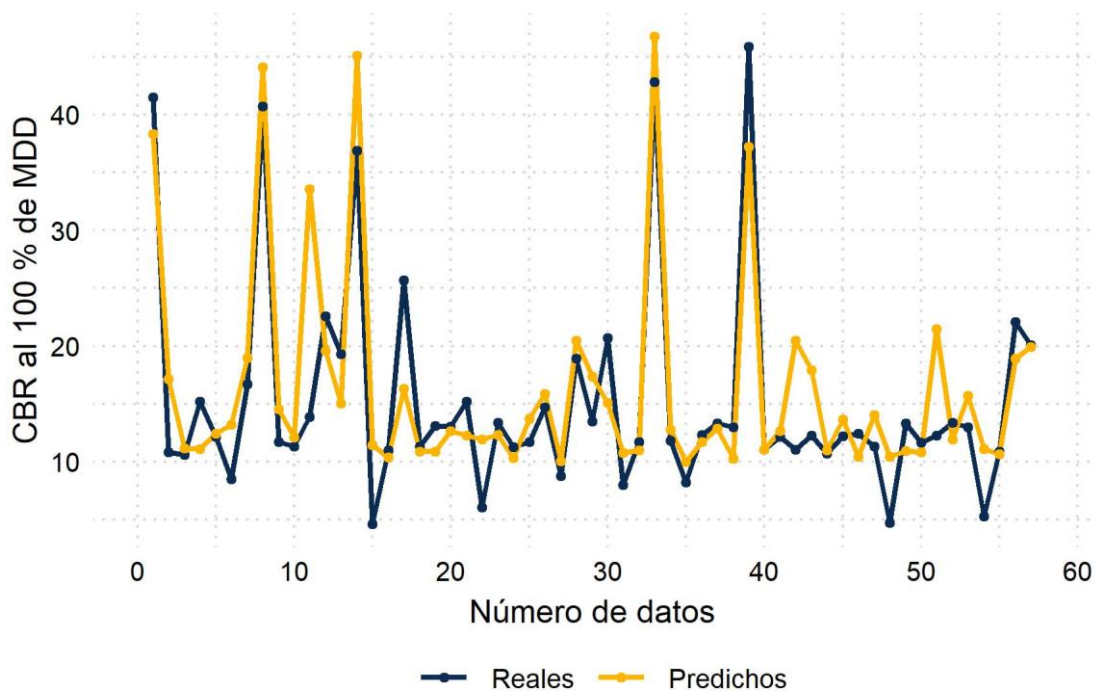
Hiperparámetros	MARK 1
Neuronas de entrada	8
Neuronas de salida	1
Capas ocultas	2
Neuronas ocultas	23
Arquitectura	8-5-18-1
Función de activación	ReLU
Tasa de aprendizaje	0.0004
Optimizador	Adam
Número de épocas	820
Tamaño de lote	16
Dropout	0.2
MSE (%)	22.13
MAE (%)	3.27

El modelo MARK 1 se volvió a entrenar utilizando todo el conjunto de entrenamiento tras aplicar la optimización bayesiana y disponer de los hiperparámetros. A continuación, se realizaron las predicciones. Así pues, la Figura

26 muestra en un gráfico de líneas que también se asemejan los valores reales de CBR al 100% de MDD a los valores predichos por el MARK 1. La figura muestra una visión general de la similitud entre ambos valores. Se puede apreciar que la red tiene un buen rendimiento. Esto se debe a que los valores predichos son similares a la variación de los ejemplos reales de CBR.

Figura 26

Predicciones realizadas por la ANN MARK 1



MARK 3

En segundo lugar, la Tabla 17 muestra los hiperparámetros del modelo de ANN MARK 3. Al igual que el modelo anterior, esta lista se generó después de aplicar la optimización bayesiana. Después de las 100 iteraciones se puede obtener que utilizando esa configuración hiperparámetros el modelo tiene un buen rendimiento para predecir el CBR al 100 % de MDD. Es importante mencionar que los resultados de MSE y MAE de la tabla son los adquiridos durante la fase de prueba, es decir,

durante las predicciones. Los errores que se han cometido están en las mismas unidades de la propiedad a predecir.

Tabla 17

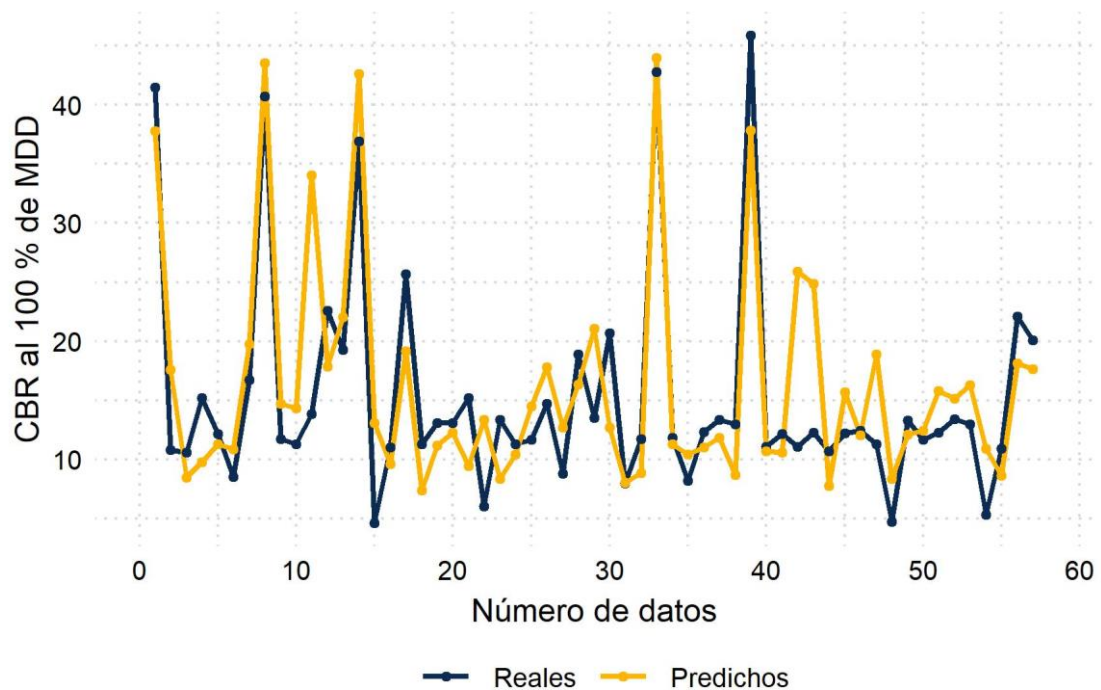
Hiperparámetros del modelo de ANN MARK 3

Hiperparámetros	MARK 3
Neuronas de entrada	10
Neuronas de salida	1
Capas ocultas	1
Neuronas ocultas	16
Arquitectura	10-16-1
Función de activación	ReLU
Tasa de aprendizaje	0.001
Optimizador	Adam
Número de épocas	235
Tamaño de lote	32
Dropout	0.3
MSE (%)	29.41
MAE (%)	4.05

El modelo MARK 3 se entrenó de nuevo utilizando todo el conjunto de entrenamiento después de aplicar la optimización bayesiana y teniendo sus hiperparámetros, se procedió a realizar las predicciones. Los valores de CBR al 100% de la MDD predichos por el MARK 3 se muestran en la Figura 27. Asimismo, a pesar de que los resultados ya mostrados anteriormente muestran que el MARK 3 tuvo un error superior al MARK 1, en la figura se puede observar que el gráfico de líneas de los valores reales y los valores predichos también son similares. Por lo tanto, utilizando el MARK 3 también predice buenos valores.

Figura 27

Predicciones realizadas por la ANN MARK 3



3.1.4.3. Mejorando el MARK 1

Tanto el MARK 1 como el MARK 3 fueron los mejores modelos para predecir el CBR al 100 % de MDD; con diferente cantidad de entradas. Ahora bien, al comparar los valores reales con los valores predichos por ambos modelos, se observó que el MARK 1 predijo el CBR al 100 % de MDD con un R de 0.8739 y el MARK 3 lo hizo con un R de 0.8313. Con estos resultados se puede ver que el rendimiento del MARK 1 fue superior al MARK 3 a la hora de realizar predicciones con datos que el modelo nunca antes había visto. Por esta razón, se hicieron cuatro modelos variando la arquitectura del MARK 1 y el número de épocas, como se ve en la Tabla 18.

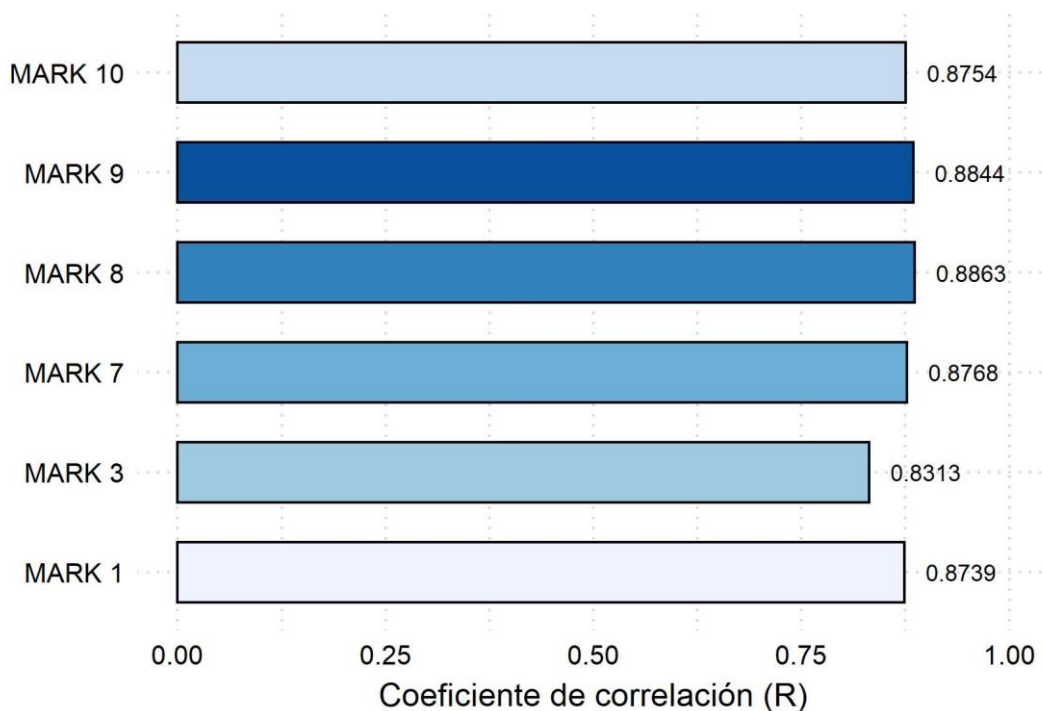
Tabla 18*Arquitectura e hiperparámetros de los modelos de ANN desarrollados para mejorar el MARK 1*

Hiperparámetros	MARK 1	MARK 7	MARK 8	MARK 9	MARK 10
Neuronas de entrada	8	8	8	8	8
Neuronas de salida	1	1	1	1	1
Capas ocultas	2	2	2	2	2
Neuronas ocultas	23	23	60	80	120
Arquitectura	8-5-18-1	8-5-18-1	8-30-30-1	8-40-40-1	8-60-60-1
Función de activación	ReLU	ReLU	ReLU	ReLU	ReLU
Tasa de aprendizaje	0.0004	0.0004	0.0004	0.0004	0.0004
Optimizador	Adam	Adam	Adam	Adam	Adam
Número de épocas	820	3000	3000	3000	3000
Tamaño de lote	16	16	16	16	16
Dropout	0.2	0.2	0.2	0.2	0.2
MSE (%)	22.13	31.90	21.80	20.42	22.27
MAE (%)	3.27	3.41	2.79	2.79	2.79

Los cuatro modelos fueron entrenados durante 3000 épocas. Asimismo, el MARK 7 tuvo la misma arquitectura del MARK 1. Por el contrario, se aumentó el número de neuronas ocultas para los modelos MARK 8, MARK 9 y MARK 10. El resto de hiperparámetros utilizados fueron los mismos del MARK 1. La Figura 28 muestra una ligera mejora en el rendimiento de los nuevos modelos de ANN desarrollados a comparación del MARK 1. Entre los modelos desarrollados destacan el MARK 8 y el MARK 9. En el gráfico de barras se puede ver que el MARK 8 tuvo un mejor rendimiento (R de 0.8863) a comparación del MARK 9 (R de 0.8844), inclusive, es mejor que el MARK 1 (R de 0.8739). Estos resultados muestran que se llegó a mejorar el MARK 1, propuesto inicialmente.

Figura 28

Rendimiento de los modelos de ANN desarrollados para mejorar el MARK 1

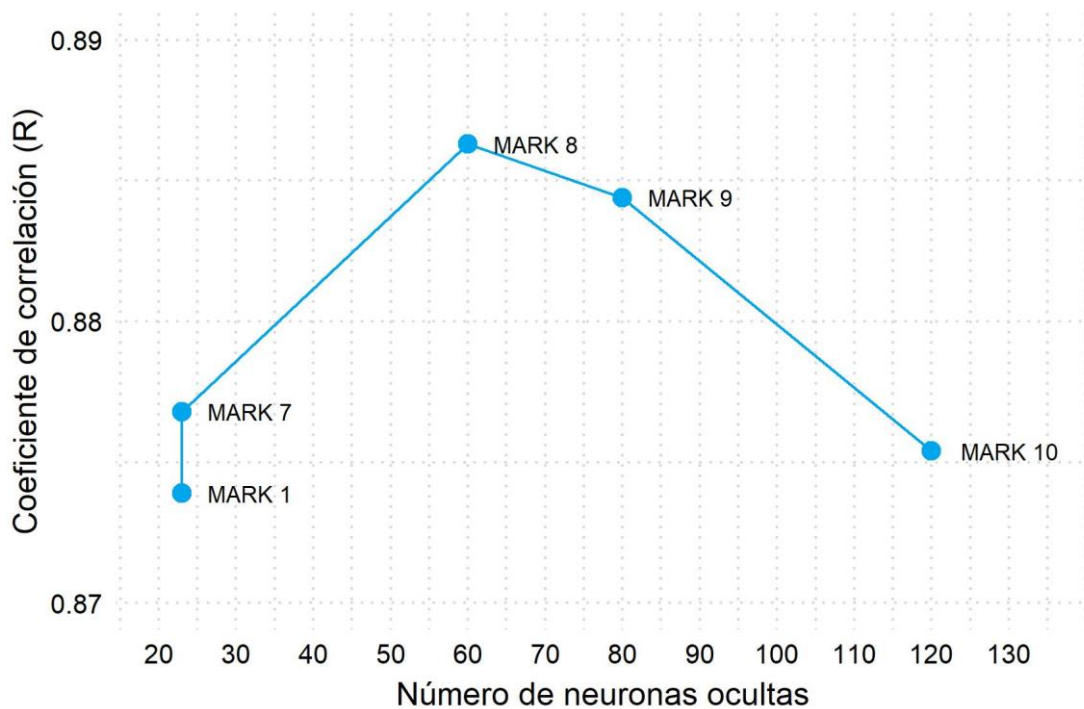


La Figura 29 muestra el rendimiento según se aumenta el número de neuronas ocultas. En la parte izquierda se puede ver que el MARK 1 y MARK 7 utilizaron el mismo número de neuronas ocultas, asimismo, se ve una mejora de la precisión con solo aumentar el número de épocas. En la figura se ve que el aumento del número de neuronas ocultas mejora la precisión; es el caso del MARK 7. Ahora

bien, es importante cuidar de que no se utilice demasiadas neuronas porque puede afectar en el rendimiento, como es el caso del MARK 9 y el MARK 10.

Figura 29

Rendimientos de los modelos desarrollados para mejorar el MARK 1 según el número de neuronas ocultas



3.1.4.4. Comparando los valores predichos y los valores reales

La Tabla 19 muestra el comportamiento de los seis modelos de ANN utilizados para predecir el CBR al 100 % de MDD. Donde el MARK 1 y MARK 3 fueron los modelos obtenidos de la búsqueda de hiperparámetros y los modelos restantes son los que se entrenaron para mejorar el MARK 1. Comparando el resultado de la correlación entre los valores reales y los valores predichos por cada uno de los modelos, se observa que el mejor modelo fue la ANN MARK 8. Por esta razón, es el modelo que se propone para predecir el CBR al 100 % de MDD:

Tabla 19*Estructura y rendimiento de los modelos de ANN desarrollados para predecir el CBR al 100 % de MDD*

ANN	ANN	Neuronas ocultas	MSE			MAE			R
			E	V	P	E	V	P	
CBR al 100 % de MDD	MARK 1	23	9.56E-03	7.51E-03	3.31E-03	6.02E-02	5.11E-02	3.99E-02	0.8739
	MARK 3	16	1.79E-02	1.66E-02	4.40E-03	7.47E-02	7.30E-02	4.95E-02	0.8313
	MARK 7	23	4.50E-03	3.69E-03	4.77E-03	4.27E-02	3.82E-02	4.17E-02	0.8768
	MARK 8	60	2.21E-03	1.27E-03	3.26E-03	2.92E-02	2.22E-02	3.42E-02	0.8863
	MARK 9	80	1.47E-03	7.85E-04	3.05E-03	2.31E-02	1.60E-02	3.41E-02	0.8844
	MARK 10	120	1.02E-03	5.41E-04	3.33E-03	1.96E-02	1.38E-02	3.41E-02	0.8754

Nota. E = entrenamiento; V = validación; P = prueba; R = coeficiente de correlación

3.1.5. Desarrollo de modelos de ANN para predecir el CBR al 95 % de MDD

Se propone un modelo para predecir el CBR al 100% de la MDD. De la misma manera, el procedimiento consistió en entrenar diferentes modelos para encontrar los hiperparámetros óptimos, hacer predicciones con el mejor modelo y comparar los valores predichos con los valores reales. Los mejores modelos fueron llamados MARK 2 y MARK 4; se seleccionaron luego de 100 iteraciones para cada uno.

3.1.5.1. Ajuste de hiperparámetros

La Tabla 20 muestra los 10 mejores modelos de ANN de un total de 100 modelos entrenados para predecir el CBR al 95% de la MDD (ver Anexo 9). Las variables de entrada utilizadas fueron H, S, G, A, F, LL, PL y PI. Se puede observar que el modelo n.º 43 fue el mejor con un MSE de 4.11E-03. En esta investigación se hace referencia a este modelo como MARK 2.

Tabla 20

Mejores 10 ANN para predecir el CBR al 95% de la MDD (8 entradas)

Ranking	Modelo	Capas ocultas	Neuronas		Learning rate	MSE
			(a)	(b)		
1	43	3	20	20	4E-05	4.11E-03
2	99	5	20	20	2E-04	4.31E-03
3	27	5	18	20	2E-05	4.35E-03
4	67	2	20	20	3E-03	4.37E-03
5	31	3	9	20	5E-05	4.49E-03
6	57	2	17	19	2E-03	4.57E-03
7	71	2	20	20	4E-03	4.64E-03
8	69	2	14	15	4E-03	4.72E-03
9	94	4	11	20	2E-04	4.83E-03
10	73	2	9	16	9E-03	4.94E-03

Nota. (a) = n.º neuronas en la primera capa oculta; (b) = n.º de neuronas en las capas ocultas restantes.

La Tabla 21 muestra los 10 mejores modelos de un total de 100 modelos entrenados para predecir el CBR al 95% de la MDD (ver Anexo 10). Se utilizó como variables de entrada H, S, G, A, F, LL, PL, PI, MDD y OMC. El modelo n.º 47 fue el modelo que mejor predijo con un MSE de 3.95E-03. En esta investigación se hace referencia a este modelo como MARK 4.

Tabla 21

Mejores 10 modelos para predecir el CBR al 95% de la MDD (10 entradas)

Ranking	Modelo	Capas ocultas	Neuronas		Learning rate	MSE
			(a)	(b)		
1	47	3	20	20	3E-03	3.95E-03
2	13	1	19	—	1E-04	4.03E-03
3	15	2	9	20	4E-05	4.06E-03
4	44	1	20	—	3E-03	4.12E-03
5	31	1	20	—	5E-03	4.14E-03
6	41	1	20	—	1E-03	4.17E-03
7	55	6	20	20	4E-03	4.25E-03
8	35	6	14	20	1E-04	4.29E-03
9	29	6	8	20	6E-05	4.29E-03
10	21	1	20	—	2E-02	4.40E-03

Nota. (a) = n.º de neuronas en la primera capa oculta; (b) = n.º de neuronas en las capas ocultas restantes.

3.1.5.2. Predicciones

Las ANN MARK 2 y MARK 4 se utilizaron para predecir los valores de CBR al 95 % de MDD, y los resultados se presentan a continuación. Para ambos modelos, el conjunto de pruebas incluye los mismos ejemplos. Asimismo, para observar la

variación de la precisión cuando el número de variables de entrenamiento aumenta o disminuye, se crearon dos modelos.

MARK 2

Para empezar, la Tabla 22 enumera los hiperparámetros utilizados por el modelo ANN MARK 2 para predecir el CBR al 95 % de MDD. Durante la búsqueda de hiperparámetros, los resultados de la optimización bayesiana tuvieron esa configuración. Esto significa que de las 100 iteraciones que utilizaron las ocho entradas, el modelo ANN MARK 2 logró los mejores resultados utilizando estos hiperparámetros. Hay que destacar que los resultados MSE y MAE de la tabla son los adquiridos durante la fase de prueba, es decir, durante las predicciones y están en las mismas unidades que el atributo predicho.

Tabla 22

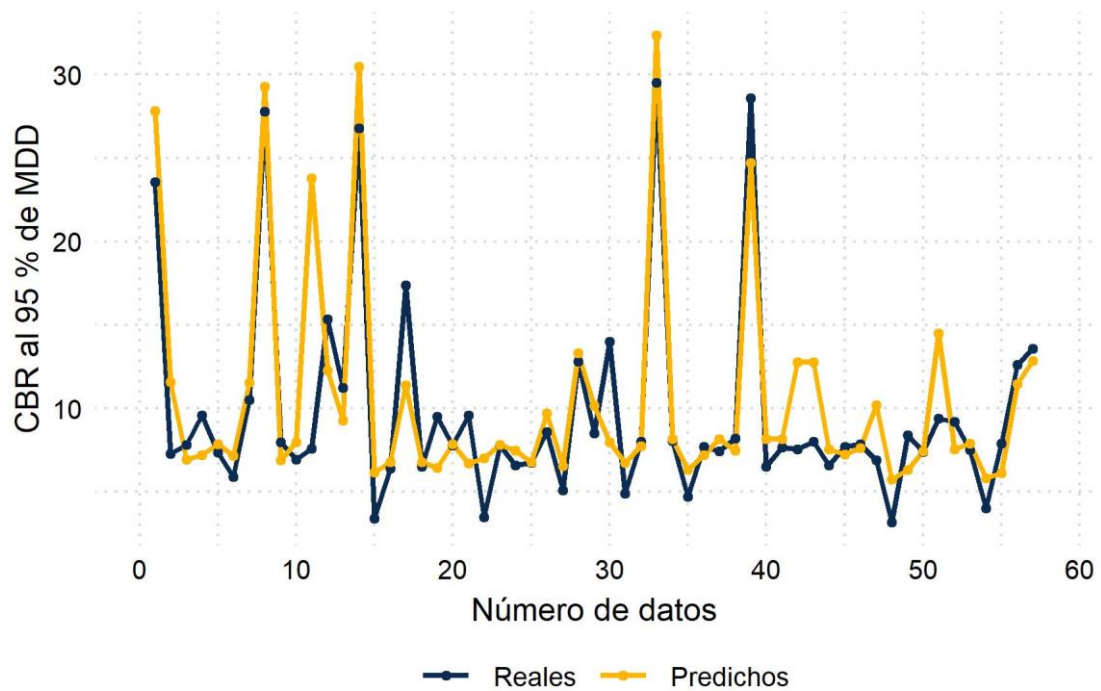
Hiperparámetros del modelo de ANN MARK 2

Hiperparámetros	MARK 2
Neuronas de entrada	8
Neuronas de salida	1
Capas ocultas	3
Neuronas ocultas	60
Arquitectura	8-20-20-20-1
Función de activación	ReLU
Tasa de aprendizaje	0.00004
Optimizador	Adam
Número de épocas	1333
Tamaño de lote	32
Dropout	0.1
MSE (%)	10.54
MAE (%)	2.12

Una vez aplicada la optimización bayesiana y teniendo los hiperparámetros, se volvió a entrenar el modelo MARK 2 utilizando todos los datos de entrenamiento. Posteriormente, se pusieron en hicieron las predicciones. A continuación, se muestran los valores de CBR al 95 % de MDD se han obtenidos con las ANN MARK 2 y MARK 4. El gráfico de líneas de la Figura 30 muestra la similitud entre los valores reales de CBR al 95% de MDD y los valores predichos por el MARK 2. Además, se puede apreciar que el gráfico de líneas de los valores predichos se asemeja al gráfico de líneas de los valores reales.

Figura 30

Predicciones realizadas por la ANN MARK 2



MARK 4

En segundo lugar, la Tabla 23 muestra los hiperparámetros del modelo ANN MARK 4. Esta lista se construyó utilizando la optimización bayesiana, al igual que el modelo anterior. Después de 100 iteraciones, queda claro que el modelo se comporta bien en la predicción de CBR al 100% de MDD utilizando esa

combinación de hiperparámetros. Cabe destacar que los valores MSE y MAE de la tabla se han obtenido durante la fase de prueba, es decir, durante las predicciones. Asimismo, los errores están en las mismas unidades que la propiedad que hay que predecir.

Tabla 23

Hiperparámetros del modelo de ANN MARK 4

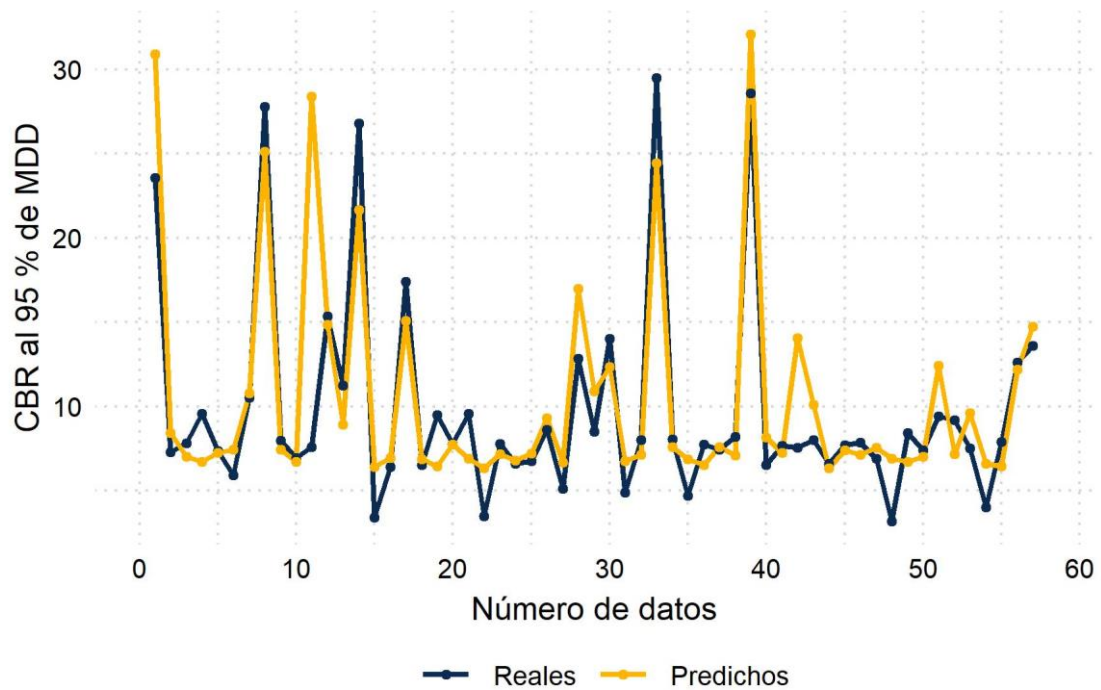
Hiperparámetros	MARK 4
Neuronas de entrada	10
Neuronas de salida	1
Capas ocultas	3
Neuronas ocultas	60
Arquitectura	10-20-20-20-1
Función de activación	ReLU
Tasa de aprendizaje	0.003
Optimizador	Adam
Número de épocas	100
Tamaño de lote	64
Dropout	0.1
MSE (%)	13.23
MAE (%)	2.11

Tras utilizar la optimización bayesiana y ajustar los hiperparámetros, se volvió a entrenar el modelo MARK 4 utilizando todos los datos de entrenamiento. Posteriormente se predijo el CBR al 95 % de MDD. El resultado de las predicciones hechas por el MARK 4 se observa en la Figura 31. Tanto el MARK 2 como el MARK 4 tuvieron el mismo número de capas ocultas diferenciándose en el número de entradas utilizadas. El conjunto de pruebas utilizado fue el mismo para ambos modelos de ANN desarrollados. En el gráfico de líneas se puede ver una existencia

de similitud entre valores reales y valores predichos. Aunque, cabe mencionar que este modelo tuvo un desempeño bajo a comparación del MARK 2.

Figura 31

Predicciones realizadas por la ANN MARK 4



3.1.5.3. Mejorando el MARK 2

El MARK 2 y MARK 4 fueron los mejores modelos para predecir el CBR al 95 % de MDD. Cuando se compararon los valores reales con los valores predichos por ambos modelos, se descubrió que MARK 2 predijo con una R de 0.8737 mientras que MARK 4 lo hizo con un R de 0.8355. Con estos resultados, se ve que el MARK 2 superó a MARK 4 a la hora de hacer predicciones. Con la finalidad de buscar un mejor rendimiento se entrenó cuatro variaciones del MARK 2 (ver Tabla 24).

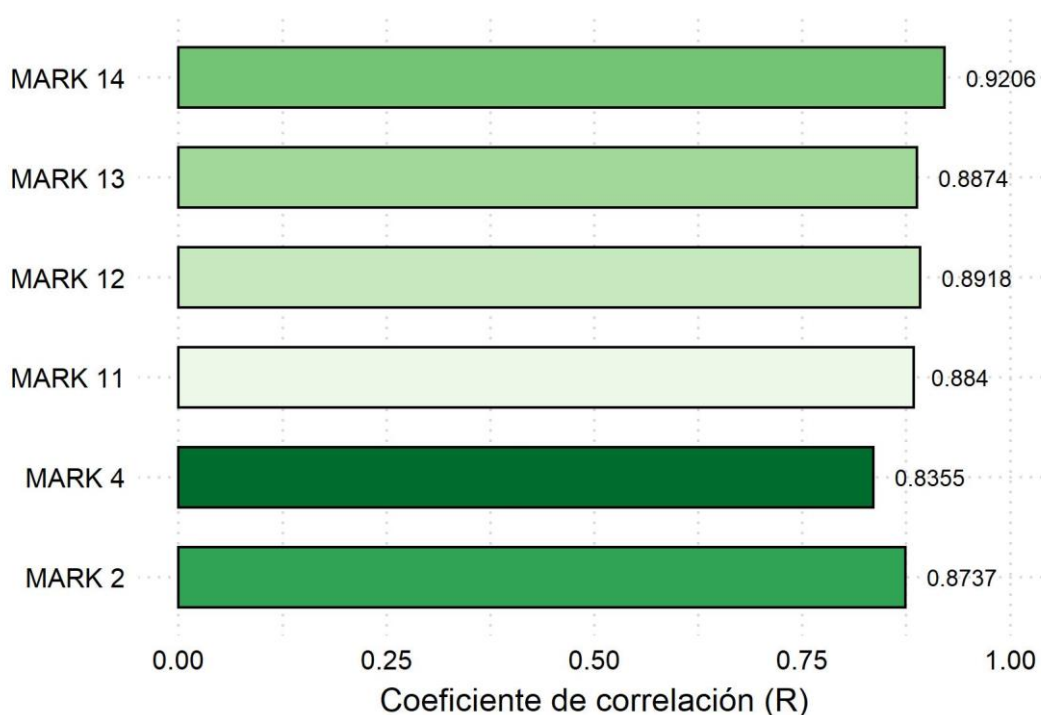
Tabla 24*Arquitectura e hiperparámetros de los modelos de ANN desarrollados para mejorar el MARK 2*

Hiperparámetros	MARK 2	MARK 11	MARK 12	MARK 13	MARK 14
Neuronas de entrada	8	8	8	8	8
Neuronas de salida	1	1	1	1	1
Capas ocultas	3	3	3	3	3
Neuronas ocultas	60	60	90	120	180
Arquitectura	8-20-20-20-1	8-20-20-20-1	8-30-30-30-1	8-40-40-40-1	8-60-60-60-1
Función de activación	ReLU	ReLU	ReLU	ReLU	ReLU
Tasa de aprendizaje	0.00004	0.00004	0.00004	0.00004	0.00004
Optimizador	Adam	Adam	Adam	Adam	Adam
Número de épocas	1333	3000	3000	3000	3000
Tamaño de lote	32	32	32	32	32
Dropout	0.1	0.1	0.1	0.1	0.1
MSE (%)	10.54	13.88	11.18	10.60	6.35
MAE (%)	2.12	2.21	2.09	1.96	1.65

Los cuatro modelos se entrenaron durante 3000 épocas. MARK 11 se construyó de la misma manera que MARK 2. En los modelos MARK 12, MARK 13 y MARK 14, sin embargo, se aumentó el número de neuronas ocultas. La Figura 32 muestra una mejora en el rendimiento de los nuevos modelos de ANN desarrollados a comparación del MARK 2. Entre los modelos desarrollados destaca el MARK 14. En el gráfico de barras se puede ver que el MARK 14 tuvo el mejor rendimiento (R igual a 0.9206) a comparación de todos los modelos de ANN desarrollados. Asimismo, el MARK 11 (R igual a 0.8840), MARK 12 (R igual a 0.8918) y MARK 13 (R igual a 0.8874) muestran un mejor rendimiento a comparación del MARK 2 (R igual a 0.8737). Estos resultados muestran que se pudo mejorar el MARK 2, propuesto inicialmente.

Figura 32

Rendimiento de los modelos de ANN desarrollados para mejorar el MARK 2

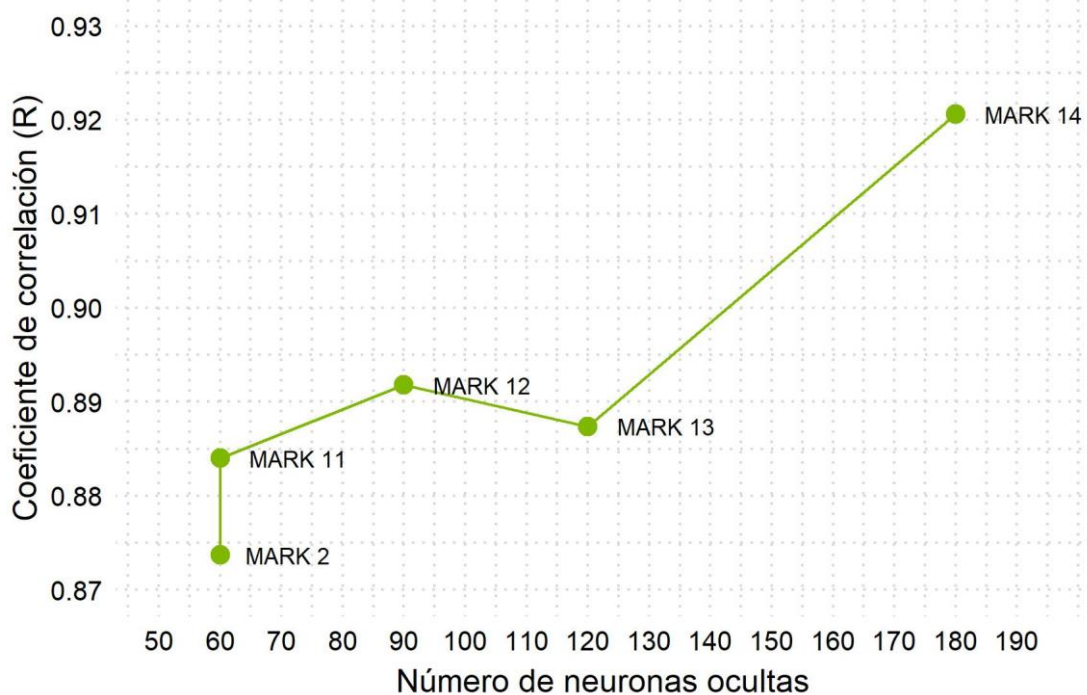


La Figura 33 muestra el rendimiento conforme aumenta el número de neuronas ocultas. En la parte izquierda se puede ver que el MARK 2 y MARK 11 utilizaron el mismo número de neuronas ocultas. También se ve una mejora de la

precisión al aumentar el número de épocas. En la figura se ve que el aumento del número de neuronas ocultas mejora la precisión, es el caso del MARK 12 y MARK 14. Sin embargo, es importante el número adecuado de neuronas ocultas porque el rendimiento puede verse afectado, como es el caso del MARK 13.

Figura 33

Rendimientos de los modelos desarrollados para mejorar el MARK 2 según el número de neuronas ocultas



3.1.5.4. Comparando los valores predichos y los valores reales

El comportamiento de los seis modelos de ANN utilizados para predecir el CBR al 95 % de MDD se muestra en la Tabla 25. Donde MARK 2 y MARK 4 son los modelos de búsqueda de hiperparámetros, y el resto de modelos son los que han sido entrenados para mejorar MARK 2. Cuando se compara la correlación entre los valores reales y los valores predichos por cada uno de los modelos, se encuentra que el mejor modelo es la ANN MARK 14. Por lo que se propone este modelo para predecir el CBR al 95 % de MDD.

Tabla 25*Estructura y rendimiento de los modelos de ANN desarrollados para predecir el CBR al 95 % de MDD*

ANN	ANN	Neuronas ocultas	MSE			MAE			R
			E	V	P	E	V	P	
CBR al 95 % de la MDD	MARK 2	60	8.52E-03	5.46E-03	2.20E-03	5.55E-02	4.10E-02	3.06E-02	0.8737
	MARK 4	60	4.11E-03	3.04E-03	2.76E-03	3.64E-02	3.13E-02	3.05E-02	0.8355
	MARK 11	60	4.36E-03	3.54E-03	2.90E-03	3.65E-02	3.10E-02	3.20E-02	0.8840
	MARK 12	90	4.22E-03	3.14E-03	2.34E-03	3.75E-02	3.03E-02	3.03E-02	0.8918
	MARK 13	120	3.14E-03	2.29E-03	2.21E-03	3.19E-02	2.50E-02	2.84E-02	0.8874
	MARK 14	180	2.64E-03	2.05E-03	1.33E-03	2.92E-02	2.38E-02	2.39E-02	0.9206

Nota. E = entrenamiento; V = validación; P = prueba; R = coeficiente de correlación

3.1.6. Desarrollo de modelos de ANN para predecir MDD

Se entrenó diversos modelos utilizando la optimización bayesiana y la validación cruzada con $k = 10$ por 3 repeticiones para encontrar los hiperparámetros óptimos, luego se hicieron predicciones con el mejor modelo y, finalmente, se comparó las predicciones realizadas por el MARK 5 con los valores reales. Todos los modelos entrenados utilizaron 8 neuronas en la capa de entrada.

3.1.6.1. Ajuste de hiperparámetros

La Tabla 26 muestra los 10 mejores modelos de ANN de un total de 100 iteraciones para predecir la MDD (ver Anexo 11). Se utilizó como variables de entrada H, S, G, A, F, LL, PL y PI. En la tabla se puede observar que el modelo n.º 10 fue el mejor modelo y, durante la búsqueda de los hiperparámetros, tuvo un MSE de $8.55E-03$. A este modelo se le llamó MARK 5.

Tabla 26

Mejores 10 modelos para predecir la MDD con 8 variables de entrada

Ranking	Modelo	Capas ocultas	Neuronas		Learning rate	MSE
			(a)	(b)		
1	10	4	15	18	3E-02	8.55E-03
2	73	6	18	20	9E-03	9.56E-03
3	63	4	18	16	6E-03	1.03E-02
4	23	5	12	9	3E-04	1.06E-02
5	9	3	13	12	3E-04	1.08E-02
6	5	1	16	—	2E-04	1.12E-02
7	22	6	20	14	1E-02	1.13E-02
8	7	1	13	—	3E-04	1.14E-02
9	82	1	20	—	5E-03	1.14E-02
10	87	2	20	20	9E-03	1.15E-02

Nota. (a) = n.º de neuronas en la primera capa oculta; (b) = n.º de neuronas en las capas ocultas restantes.

3.1.6.2. Predicciones

La Tabla 27 muestra los hiperparámetros del modelo ANN MARK 5. Esta lista se creó utilizando la optimización bayesiana. Después de 100 iteraciones, el modelo tiene buen rendimiento en la predicción de MDD utilizando ese conjunto de hiperparámetros. Los valores MSE y MAE que se muestran en la tabla se recogieron durante la fase de pruebas.

Tabla 27

Arquitectura e hiperparámetros de los mejores modelos aplicando optimización bayesiana

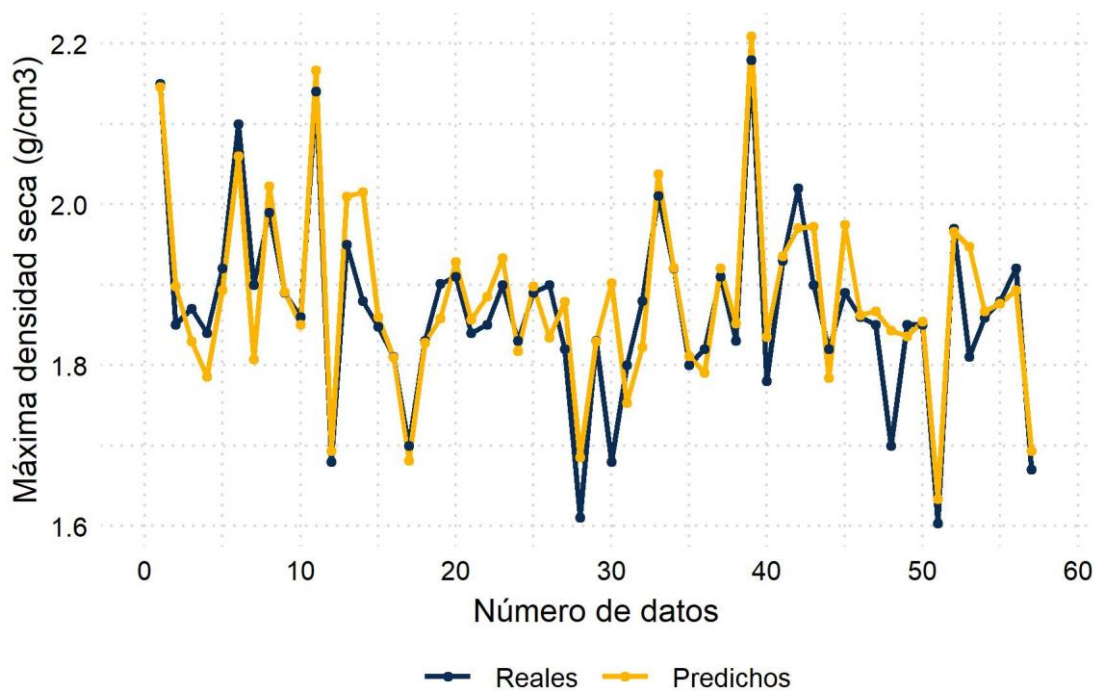
Hiperparámetros	MARK 5
Neuronas de entrada	8
Neuronas de salida	1
Capas ocultas	4
Neuronas ocultas	69
Arquitectura	8-15-18-18-18-1
Función de activación	ReLU
Tasa de aprendizaje	0.03
Optimizador	Adam
Número de épocas	1218
Tamaño de lote	64
Dropout	0.0
MSE (g/cm ³)	0.003
MAE (g/cm ³)	0.038

El modelo MARK 5 se volvió a entrenar utilizando todos los datos de entrenamiento una vez que se realizó la optimización bayesiana y se tuvieron los

hiperparámetros. A continuación, se realizaron las predicciones de esta propiedad. La Figura 34 muestra el grado de proximidad entre los valores reales de la MDD y los valores predichos por MARK 5.

Figura 34

Predicciones realizadas por la ANN MARK 5



3.1.6.3. Mejorando el MARK 5

Se buscó mejorar el rendimiento para predecir la MDD. Para ello, se decidió realizar algunas variaciones al modelo de ANN desarrollado; el MARK 5. El procedimiento consistió en entrenar 4 modelos más de ANN, realizar predicciones de la MDD y comparar los valores predichos por cada modelo con los valores reales (ver Tabla 28).

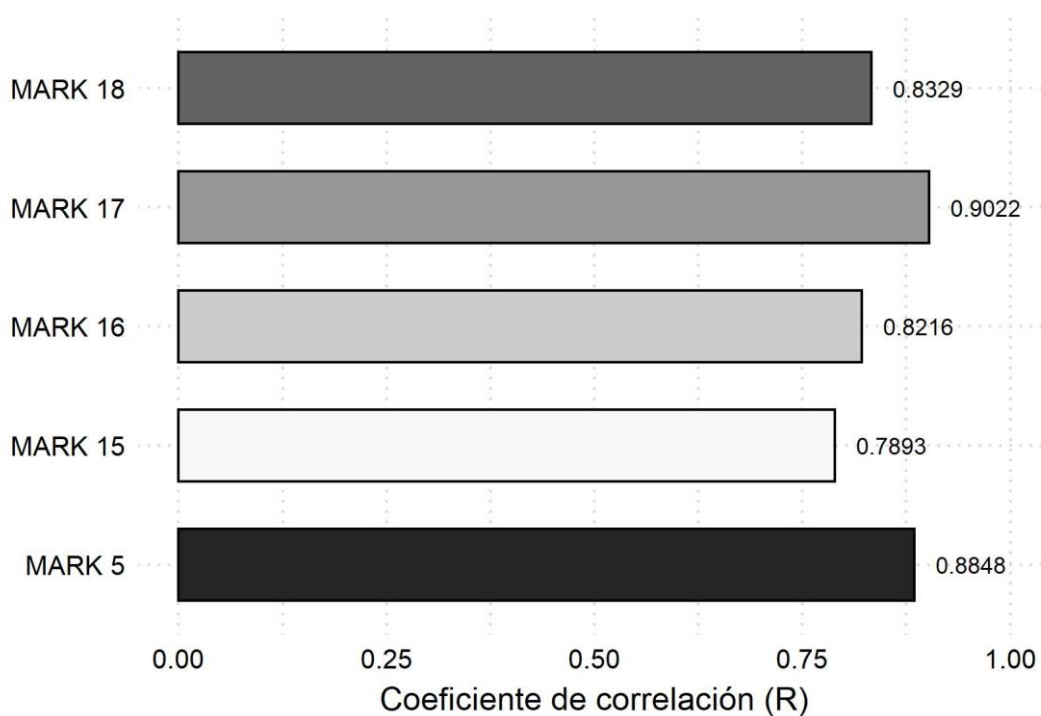
Tabla 28*Arquitectura e hiperparámetros de los modelos de ANN desarrollados para mejorar el MARK 5*

Hiperparámetros	MARK 5	MARK 15	MARK 16	MARK 17	MARK 18
Neuronas de entrada	8	8	8	8	8
Neuronas de salida	1	1	1	1	1
Capas ocultas	4	4	4	4	4
Neuronas ocultas	69	69	120	160	240
Arquitectura	8-15-18-18-18-1	8-15-18-18-18-1	8-30-30-30-30-1	8-40-40-40-40-1	8-60-60-60-60-1
Función de activación	ReLU	ReLU	ReLU	ReLU	ReLU
Tasa de aprendizaje	0.03	0.03	0.03	0.03	0.03
Optimizador	Adam	Adam	Adam	Adam	Adam
Número de épocas	1218	3000	3000	3000	3000
Tamaño de lote	64	64	64	64	64
Dropout	0.0	0.0	0.0	0.0	0.0
MSE (g/cm ³)	0.003	0.005	0.005	0.003	0.005
MAE (g/cm ³)	0.038	0.050	0.046	0.035	0.043

La Figura 35 muestra una disminución más que un aumento en el rendimiento de los modelos de ANN a comparación del MARK 5. Entre los modelos desarrollados el único que destaca es el MARK 17. En el gráfico de barras se puede ver que el MARK 17 tuvo el mejor rendimiento (R igual a 0.9022) a comparación de todos los modelos de ANN desarrollados. Asimismo, el MARK 15, MARK 16 y MARK 18 muestran un rendimiento por debajo del MARK 5. A pesar de ello, los resultados muestran que se pudo mejorar el MARK 5, propuesto en un inicio.

Figura 35

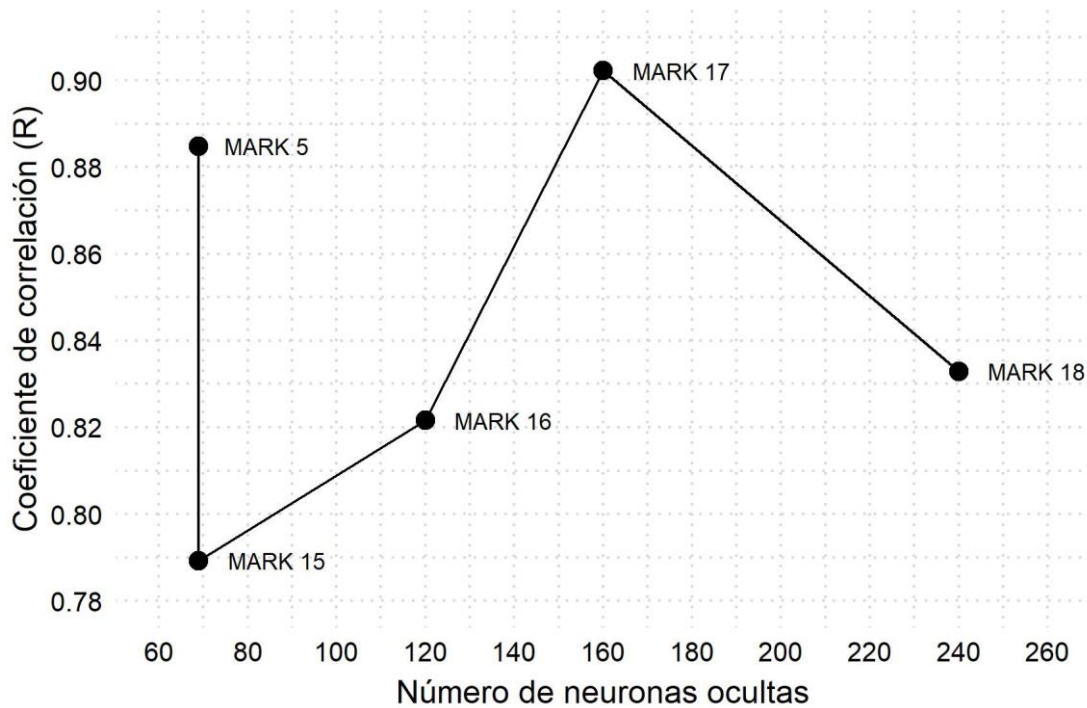
Rendimiento de los modelos de ANN desarrollados para mejorar el MARK 5



La Figura 36 muestra el rendimiento conforme se aumenta el número de neuronas ocultas. En la parte izquierda se puede ver que el MARK 5 y MARK 15 utilizaron el mismo número de neuronas ocultas. También se ve una pérdida de la precisión al aumentar el número de épocas. En la figura se ve que el aumento del número de neuronas ocultas puede mejorar la precisión de una ANN. Sin embargo, el número de neuronas ocultas erróneo (MARK 16), así como, el exceso de estas, pueden afectar en el rendimiento, como es el caso del MARK 18.

Figura 36

Rendimientos de los modelos desarrollados para mejorar el MARK 5 según el número de neuronas ocultas



3.1.6.4. Comparando los valores predichos y los valores reales

Como ya se mencionó anteriormente, de los 100 modelos entrenados el MARK 5 fue el mejor modelo para predecir el MDD. Para evaluar si el modelo era capaz de generalizar su conocimiento se comparó los valores reales con los valores predichos y se observó que el MARK 17 tuvo un buen rendimiento a la hora de realizar predicciones con datos que el modelo nunca antes había visto (ver Tabla 29). Por esta razón se propone el modelo de ANN MARK 17 para predecir MDD.

Tabla 29*Estructura y rendimiento de los modelos de ANN desarrollados para predecir MDD*

ANN	ANN	Neuronas ocultas	MSE			MAE			R
			E	V	P	E	V	P	
MDD	MARK 5	69	4.49E-03	4.26E-03	7.42E-03	4.65E-02	4.59E-02	5.80E-02	0.8848
	MARK 15	69	5.86E-03	5.68E-03	1.25E-02	5.53E-02	5.47E-02	7.69E-02	0.7893
	MARK 16	120	1.45E-03	1.35E-03	1.29E-02	2.67E-02	2.61E-02	7.05E-02	0.8216
	MARK 17	160	1.74E-03	1.63E-03	7.31E-03	2.87E-02	2.81E-02	5.40E-02	0.9022
	MARK 18	240	3.67E-03	3.49E-03	1.12E-02	4.18E-02	4.13E-02	6.55E-02	0.8329

Nota. E = entrenamiento; V = validación; P = prueba; R = coeficiente de correlación

3.1.7. Desarrollo de modelos de ANN para predecir OMC

Para encontrar los parámetros óptimos se utilizó la optimización bayesiana y la validación cruzada con $K = 10$ durante 3 repeticiones, se seleccionó el mejor modelo para hacer predicciones con un conjunto de prueba que el modelo no conocía y se comparó los valores reales con los valores predichos. Cabe mencionar que los modelos entrenados utilizaron 8 neuronas en la capa de entrada.

3.1.7.1. Ajuste de hiperparámetros

La Tabla 30 muestra los 10 mejores modelos de ANN de un total de 100 modelos entrenados para predecir el OMC (ver Anexo 12). Las variables de entrada utilizadas en el modelo fueron H, S, G, A, F, LL, PL y PI. Asimismo, se puede observar que el modelo n.º 58 fue el mejor y tuvo un MSE de $1.49E-02$. En esta investigación se hará referencia a este modelo como MARK 6.

Tabla 30

Mejores 10 modelos para predecir el OMC con 8 variables de entrada

Ranking	Modelo	Capas ocultas	Neuronas		Learning rate	MSE
			(a)	(b)		
1	58	2	19	20	7E-04	1.49E-02
2	30	4	20	20	5E-05	1.51E-02
3	98	4	18	18	8E-04	1.54E-02
4	66	2	9	19	2E-03	1.55E-02
5	83	2	20	20	4E-04	1.56E-02
6	1	3	19	15	2E-03	1.56E-02
7	55	3	20	20	7E-04	1.61E-02
8	2	4	20	20	3E-05	1.64E-02
9	28	4	20	20	4E-04	1.64E-02
10	51	3	13	20	7E-04	1.66E-02

Nota. (a) = n.º de neuronas en la primera capa oculta; (b) = n.º de neuronas en las capas ocultas restantes.

3.1.7.2. Predicción

Los hiperparámetros utilizados por el modelo ANN MARK 6 para predecir OMC se enumeran en la Tabla 31. Los resultados de la optimización bayesiana produjeron esa configuración durante la búsqueda de hiperparámetros. Esto indica que, con estos hiperparámetros, el modelo ANN MARK 6 produjo los mejores resultados de 100 iteraciones utilizando las ocho entradas. Los resultados MSE y MAE de la tabla se obtuvieron durante la fase de prueba, es decir, durante las predicciones, y están en las mismas unidades que OMC.

Tabla 31

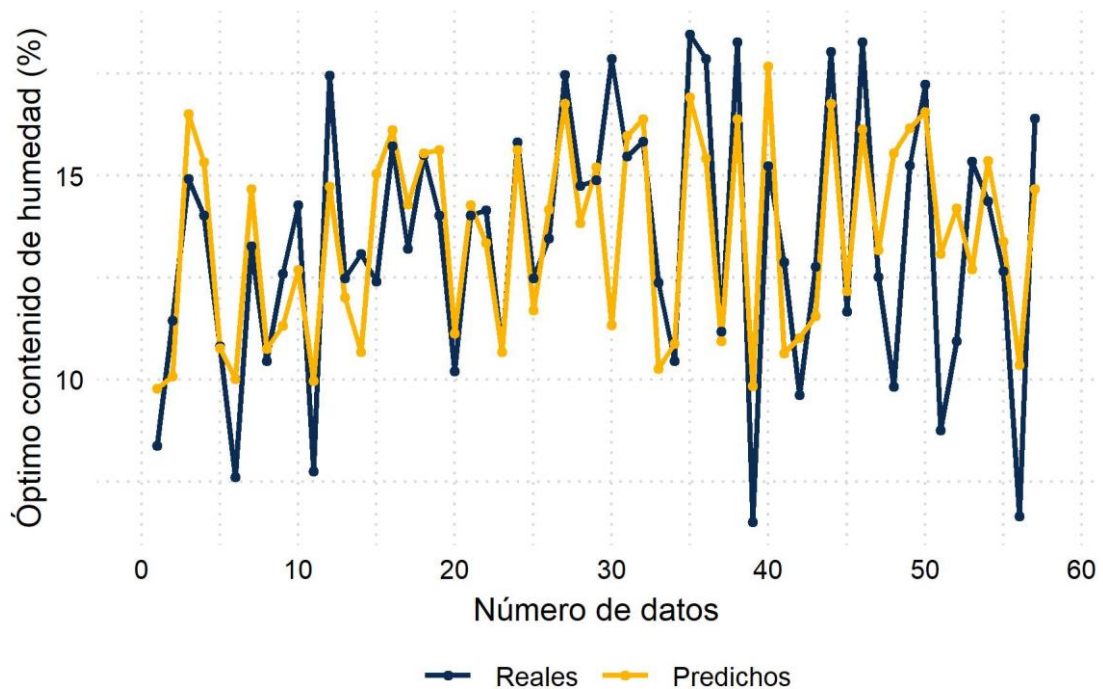
Arquitectura e hiperparámetros de los mejores modelos aplicando optimización bayesiana

Hiperparámetros	MARK 6
Neuronas de entrada	8
Neuronas de salida	1
Capas ocultas	2
Neuronas ocultas	39
Arquitectura	8-19-20-1
Función de activación	ReLU
Tasa de aprendizaje	0.0007
Optimizador	Adam
Número de épocas	188
Tamaño de lote	32
Dropout	0.2
MSE (%)	4.11
MAE (%)	1.54

El modelo MARK 6 se volvió a entrenar utilizando todos los datos de entrenamiento una vez que se realizó la optimización bayesiana y se tuvieron los hiperparámetros. La Figura 37 muestra el resultado de las predicciones realizadas por la ANN MARK 6. En la figura se observa que el gráfico de líneas de los valores reales del OMC se asemeja al gráfico de los valores predichos por la ANN

Figura 37

Predicciones realizadas por la ANN MARK 6



3.1.7.3. Mejorando el MARK 6

De los resultados mostrados en la parte anterior, el MARK 6 tuvo un bajo rendimiento a comparación de los demás. Por esta razón, se decidió hacer más modelos en busca de mejorar la precisión al predecir el OMC. El procedimiento consistió en entrenar 4 modelos más de ANN, realizar predicciones del OMC y comparar los valores predichos por cada modelo con los valores reales, como se muestra en la Tabla 32.

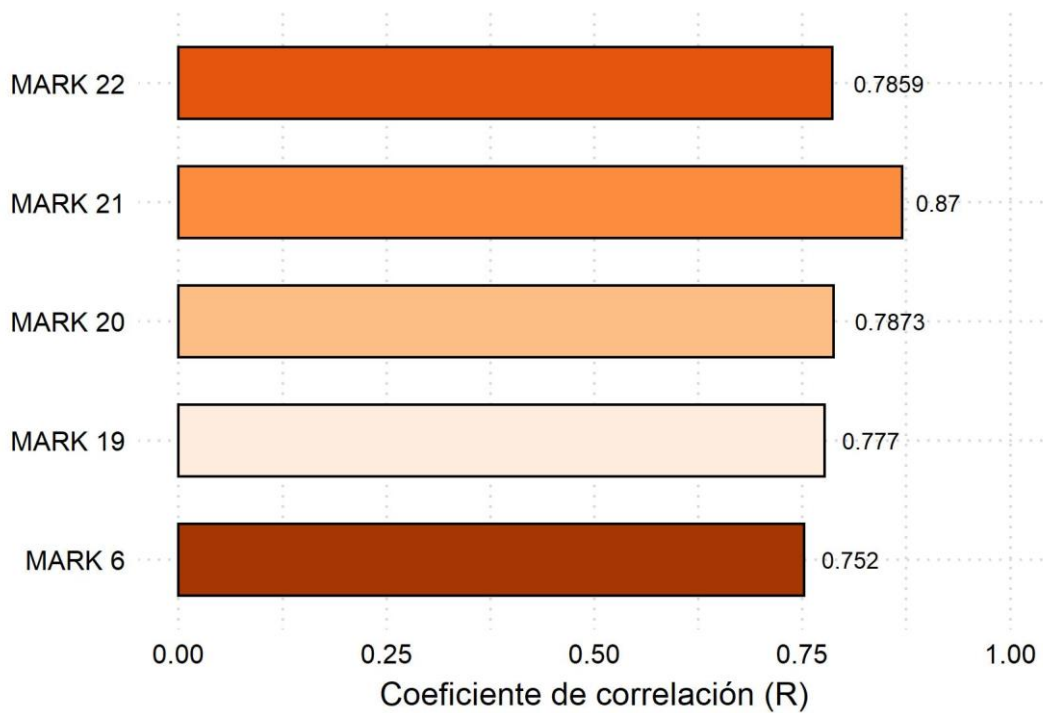
Tabla 32*Arquitectura e hiperparámetros de los modelos de ANN desarrollados para mejorar el MARK 6*

Hiperparámetros	MARK 6	MARK 19	MARK 20	MARK 21	MARK 22
Neuronas de entrada	8	8	8	8	8
Neuronas de salida	1	1	1	1	1
Capas ocultas	2	2	2	2	2
Neuronas ocultas	39	39	60	80	120
Arquitectura	8-19-20-1	8-19-20-1	8-30-30-1	8-40-40-1	8-60-60-1
Función de activación	ReLU	ReLU	ReLU	ReLU	ReLU
Tasa de aprendizaje	0.0007	0.0007	0.0007	0.0007	0.0007
Optimizador	Adam	Adam	Adam	Adam	Adam
Número de épocas	188	3000	3000	3000	3000
Tamaño de lote	32	32	32	32	32
Dropout	0.2	0.2	0.2	0.2	0.2
MSE (%)	4.11	3.86	3.89	2.42	4.16
MAE (%)	1.54	1.40	1.38	1.13	1.44

La Figura 38 muestra un aumento del rendimiento de los modelos de ANN a comparación del MARK 6. Entre los modelos desarrollados el que más destaca es el MARK 21. En el gráfico de barras se puede ver que el MARK 21 tuvo el mejor rendimiento a comparación de todos los modelos de ANN desarrollados. Asimismo, el MARK 19, MARK 20 y MARK 22 muestran un rendimiento por encima del MARK 6. Los resultados muestran que se pudo mejorar el MARK 6, propuesto en un inicio.

Figura 38

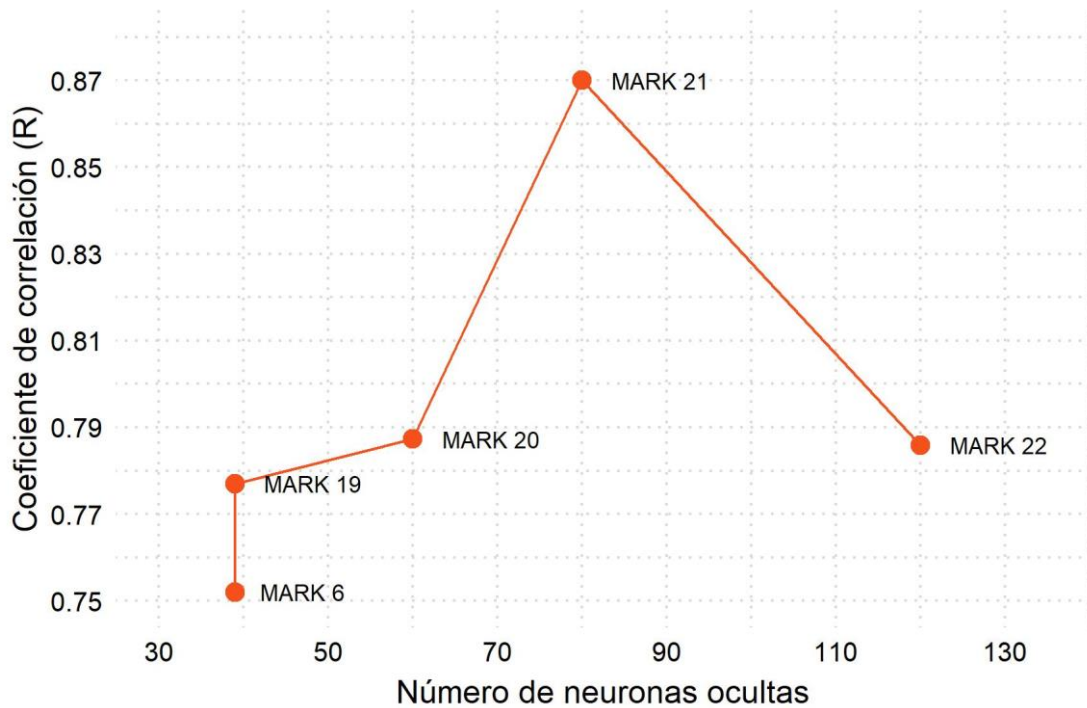
Rendimiento de los modelos de ANN desarrollados para mejorar el MARK 6



La Figura 39 muestra el rendimiento conforme se aumenta el número de neuronas ocultas. En la parte izquierda de la figura se puede ver que el MARK 6 y MARK 19 utilizaron el mismo número de neuronas ocultas. También se ve que la precisión mejora al aumentar el número de épocas. En la figura se observa que el aumento del número de neuronas ocultas mejora la precisión, como el MARK 20 y MARK 21. Sin embargo, una gran cantidad de neuronas en cada capa oculta, puede afectar en la precisión del modelo, como es el caso del MARK 22.

Figura 39

Rendimientos de los modelos desarrollados para mejorar el MARK 6 según el número de neuronas ocultas



3.1.7.4. Comparando los valores predichos y los valores reales

Como ya se mencionó anteriormente, de 100 iteraciones MARK 6 fue el mejor modelo para predecir OMC. Para evaluar si el modelo era capaz de generalizar su conocimiento se comparó los valores reales con los valores predichos y se observó que el MARK 21 tuvo el mejor comportamiento para predecir OMC con datos que el modelo nunca había visto (ver Tabla 33). Por lo que es el modelo que se propone para predecir el OMC.

Tabla 33*Estructura y rendimiento de los modelos de ANN desarrollados*

ANN	ANN	Neuronas ocultas	MSE			MAE			R
			E	V	P	E	V	P	
OMC	MARK 6	39	2.91E-02	2.08E-02	1.51E-02	1.23E-01	1.01E-01	9.37E-02	0.7520
	MARK 19	39	1.01E-02	7.79E-03	1.42E-02	7.21E-02	6.13E-02	8.51E-02	0.7770
	MARK 20	60	6.76E-03	4.96E-03	1.43E-02	5.86E-02	4.87E-02	8.35E-02	0.7873
	MARK 21	80	5.98E-03	4.47E-03	8.92E-03	5.33E-02	4.38E-02	6.83E-02	0.8700
	MARK 22	120	5.01E-03	3.73E-03	1.53E-02	4.83E-02	3.94E-02	8.71E-02	0.7859

Nota. E = entrenamiento; V = validación; P = prueba; R = coeficiente de correlación

3.2. Discusión de resultados

El propósito de este estudio era predecir CBR al 100 % de MDD, CBR al 95 % de MDD, MDD y OMC por medio del desarrollo de modelos de ANN. En esta parte de la investigación se discuten los principales hallazgos en relación con la literatura sobre el desarrollo de modelos de ANN para predecir el CBR y los parámetros de compactación. Asimismo, la determinación óptima de hiperparámetros como el número de neuronas ocultas y el número de épocas.

La teoría sobre el comportamiento del suelo señala que son diversos los factores que afectan en sus propiedades mecánicas, entre ellos están la composición química y la formación de los suelos. Asimismo, el comportamiento del suelo dependerá de las características físicas que este posea. Todos estos factores hacen que predecir las propiedades mecánicas del suelo se torne complicado. Sin embargo, el avance de la ciencia y la tecnología, como la aplicación de las ANN, ayuda a tener herramientas que ayuden a predecir las propiedades del suelo en poco tiempo.

3.2.1. Aplicación de modelos de ANN para predecir las propiedades del suelo

Respecto a la hipótesis planteada en esta investigación, según los resultados obtenidos pudo llegar a cumplirse, puesto que la precisión de los modelos desarrollados fue mayor al 80 %. Además, la conclusión de esta investigación de que los modelos de ANN pueden predecir el CBR y los parámetros de compactación del suelo coincide con la literatura que indica que los modelos de ANN son muy exactas y se pueden aplicar para predecir el CBR y los parámetros de compactación (Alam *et al.*, 2020; Albusoda *et al.*, 2020; Erzin & Turkoz, 2016; Fikret Kurnaz & Kaya, 2019; Ghorbani & Hasanzadehshooiili, 2018; Rajakumar & Reddy babu, 2020; Suthar & Aggarwal, 2018).

Mediante el coeficiente de correlación (R) se demostró que se puede predecir el CBR y los parámetros de compactación del suelo aplicando ANN. Los resultados al predecir CBR al 100 % de MDD (R de 0.8863) y CBR al 95 % (R de 0.9206) fueron similares a comparación del CBR al 100 % de MDD y CBR al 95 %

de MDD de otra investigación. De igual manera, MDD (R de 0.9022) y OMC (R de 0.8700) comparado a MDD y OMC de otra investigación (Manrique Dávila, 2019).

3.2.2. Predicción del CBR y los parámetros de compactación según las propiedades físicas del suelo

La predicción del CBR al 100 % y CBR al 95 % de MDD, MDD y OMC del suelo fue posible debido a que se usaron propiedades que influyen en el comportamiento del suelo como H, S, G, A, F, LL, PL y PI (Ardakani & Kordnaeij, 2019; Tenpe & Patel, 2020). En este estudio se pudo encontrar que el aumento del número de neuronas de entradas (de 8 a 10) disminuyó la precisión de los modelos de ANN para predecir el CBR al 100 % de MDD.

3.2.3. El número de neuronas y la precisión

En este estudio se evaluó la precisión de los modelos de ANN al cambiar la cantidad de neuronas en cada capa oculta. Se encontró que el aumento de la cantidad de neuronas ocultas mejora la exactitud de las predicciones de los modelos de ANN desarrollados. Sin embargo, también se encontró que una mayor cantidad de neuronas ocultas puede causar un efecto contrario, es decir que la precisión del modelo se puede ver afectada.

Velasco *et al.* (2019) menciona en su investigación la fórmula de Stathakis' para tener la cantidad aproximada de neuronas en cada capa oculta. Está bien para tener el número de neuronas inicial a utilizar cuando se está comenzando a desarrollar un modelo de ANN. Esta fórmula fue aplicada en esta investigación y se entrenaron los modelos. Pero, según los resultados obtenidos, se consiguió una mejor precisión de los modelos con un número de neuronas por cada capa oculta mucho mayor al calculado por la fórmula.

3.2.4. El tiempo de entrenamiento afecta en la precisión

En este estudio se pudo verificar que, si se aumenta el tiempo de entrenamiento del modelo de ANN a desarrollar, se mejora la precisión (Diaz *et al.*, 2017; Kim, 2019). Aunque, también se pudo encontrar que un exceso de entrenamiento puede

hacer que el modelo aprenda más de lo que debería, llegando a tener un mal comportamiento con datos nuevo.

Con esto se puede ver que no hay una forma mágica que permite obtener la cantidad de neuronas a utilizar en cada capa oculta o el número de épocas adecuado. El desarrollo de modelos de ANN consiste más en un procedimiento de prueba y error. Viendo como varía el comportamiento de los modelos según las características del conjunto de datos. Todo esto, claro, estando de la mano con las recomendaciones hechas en las teorías.

IV. CONCLUSIONES Y RECOMENDACIONES

4.1. Conclusiones

En este trabajo se desarrolló modelos de ANN para predecir las propiedades mecánicas del suelo de la región Lambayeque. Se realizó a partir de propiedades físicas del suelo fáciles de medir, con el fin de facilitar la obtención de los resultados, que estos sean precisos y, además, se obtengan en el menor tiempo posible. Las conclusiones a las que se llegó, de acuerdo a los resultados obtenidos en cada etapa de este estudio y siguiendo los objetivos, se describen de la siguiente manera:

- Se definió las propiedades mecánicas del suelo a predecir con el desarrollo de modelos de ANN. Lo importante de tener definido las propiedades que se pueden predecir es identificar las propiedades que pueden utilizarse como variables de entrada para la predicción de dicha propiedad. Para ello, ayudó el análisis de la literatura, el número de atributos disponible por cada muestra de suelo y el análisis de correlación entre variables.
- Se determinó los hiperparámetros óptimos de cada ANN mediante el lenguaje de programación de Python utilizando Spyder. Lo importante de la determinación de los hiperparámetros fue encontrar la ANN con mayor capacidad de generalizar. Lo que ayudó a identificar los hiperparámetros óptimos fue el MSE porque muestra que tan bien un modelo redujo el error y cómo se comportó con respecto a otro. Lo difícil en la identificación de los hiperparámetros fue realizar el código en Spyder porque hubo diversas técnicas aplicables.
- Se demostró que cada ANN aprendió correctamente comparando los valores reales con los valores predichos por el modelo. Lo importante fue obtener valores de R altos para MDD, OMC, CBR al 95% de MDD y CBR al 100% de MDD, ya que permite demostrar la elevada capacidad de las ANN para predecir propiedades del suelo en base a propiedades fáciles de medir, aun cuando se tiene pocos ejemplos y son varios los elementos

que influyen en el comportamiento del suelo. Además, los modelos propuestos pueden usarse en proyectos reales cuando se quiere obtener estas propiedades de manera rápida.

4.2. Recomendaciones

Los modelos de ANN propuestos en esta investigación para predecir MDD, OMC, CBR al 95% de MDD y CBR al 100% de MDD han sido desarrollados con un conjunto de datos pequeño, llegando a obtener buenos resultados. Estos modelos pueden mejorarse y la principal recomendación es aumentar la cantidad de ejemplos del conjunto de datos.

Se recomienda que los resultados de los ensayos del Laboratorio de Mecánica de Suelos y Rocas de la escuela de Ingeniería Civil de la Universidad Señor de Sipán se guarden en una base de datos, de esta manera se puede llegar a tener una gran cantidad de ejemplos que pueden ser utilizados para desarrollar modelos de ANN en futuras investigaciones.

En esta investigación se ha demostrado que las ANN desarrolladas presentan un buen rendimiento para predecir propiedades del suelo. Es importante ampliar este conocimiento y aplicarlo en otras ramas de la ingeniería civil. Una de las recomendaciones es realizar investigaciones sobre la aplicación de ANN para predecir las propiedades del concreto, diseño de estructuras y otras propiedades del suelo.

Referencias

- Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A. E., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, 4(11). <https://doi.org/10.1016/j.heliyon.2018.e00938>
- Al-Busultan, S., Aswed, G. K., Almuhanha, R. R. A., & Rasheed, S. E. (2020). Application of Artificial Neural Networks in Predicting Subbase CBR Values Using Soil Indices Data. *IOP Conference Series: Materials Science and Engineering*, 671(1). <https://doi.org/10.1088/1757-899X/671/1/012106>
- Alam, S. K., Mondal, A., & Shiuly, A. (2020). Prediction of CBR Value of Fine Grained Soils of Bengal Basin by Genetic Expression Programming, Artificial Neural Network and Krigging Method. *Journal of the Geological Society of India*, 95(2), 190–196. <https://doi.org/10.1007/s12594-020-1409-0>
- Albusoda, B. S., Al-Hamdani, D. A., & Abbas, M. F. (2020). Dry Density Based on Soil Index Properties by Using Expert System. *Key Engineering Materials*, 857, 266–272. <https://doi.org/10.4028/www.scientific.net/KEM.857.266>
- Ali, Z., Hussain, I., Faisal, M., Nazir, H. M., Hussain, T., Shad, M. Y., Mohamd Shoukry, A., & Hussain Gani, S. (2017). Forecasting Drought Using Multilayer Perceptron Artificial Neural Network Model. *Advances in Meteorology*, 2017, 1–9. <https://doi.org/10.1155/2017/5681308>
- Amirabadi, M. A., Kahaei, M. H., & Nezamalhoseini, S. A. (2020). Novel suboptimal approaches for hyperparameter tuning of deep neural network [under the shelf of optical communication]. *Physical Communication*, 41, Article 101057. <https://doi.org/10.1016/j.phycom.2020.101057>
- Ardakani, A., & Kordnaeij, A. (2019). Soil compaction parameters prediction using GMDH -type neural network and genetic algorithm. *European Journal of Environmental and Civil Engineering*, 23(4), 449–462. <https://doi.org/10.1080/19648189.2017.1304269>
- Balles, L., Romero, J., & Hennig, P. (2017, December 15). Coupling adaptive batch sizes with learning rates. *Uncertainty in Artificial Intelligence - Proceedings of the 33rd Conference, UAI 2017*. <http://arxiv.org/abs/1612.05086>
- Belmont, I. (1979). *Principios éticos y directrices para la protección de sujetos humanos de investigación*.

- Benitez, R. (2014). *Inteligencia artificial avanzada*. Editorial UOC.
<https://elibro.net/es/lc/bibsipan/titulos/57582>
- Bogado, G. O., Pintos, N. A., Reinert, H. O., & Bressan, D. A. (2017). Correlación entre parámetros de compactación y propiedades geotécnicas en suelos tropicales misioneros. *Revista de Geología Aplicada a La Ingeniería y Al Ambiente*, 39, 19–26.
- Braja M. Das, & Nagaratnam Sivakugan. (2019). *Principles of Foundation Engineering* (9th ed.). Cengage Learning.
- Brochu, E., Cora, V. M., & de Freitas, N. (2010). A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. *RI*. <http://arxiv.org/abs/1012.2599>
- Carlson, D., Cevher, V., & Carin, L. (2015). Stochastic spectral descent for restricted boltzmann machines. *Journal of Machine Learning Research*, 38, 111–119.
- Carrillo gil, A. (2017). Comportamiento del suelo tropical peruano. *Perfiles de Ingeniería*, 12(12), 27–35.
https://doi.org/10.31381/perfiles_ingenieria.v12i12.809
- Casas Roma, J., Bosch Rue, A., & Lozano Bagen, T. (2019). *Deep learning: principios y fundamentos*. Editorial UOC.
<https://elibro.net/es/lc/bibsipan/titulos/126167>
- Cruz del Castillo, C., & Socorro Olivares, O. (2014). *Metodología de la investigacion*. Grupo Editorial Patria. <https://elibro.net/es/lc/bibsipan/titulos/39410>
- Dauphin, Y. N., de Vries, H., & Bengio, Y. (2015). Equilibrated adaptive learning rates for non-convex optimization. *Advances in Neural Information Processing Systems*, 2015-Janua, 1504–1512.
<http://arxiv.org/abs/1502.04390>
- Diaz, G. I., Fokoue-Nkoutche, A., Nannicini, G., & Samulowitz, H. (2017). An effective algorithm for hyperparameter optimization of neural networks. *IBM Journal of Research and Development*, 61(4/5).
<https://doi.org/10.1147/JRD.2017.2709578>
- Dinamarca, A. (2018). *Aprendizaje y Análisis de Redes Neuronales Artificiales Profundas* [Tesina de grado, Universidad Nacional de Cuyo].
<https://bdigital.uncu.edu.ar/13989>
- Ecer, F., Ardabili, S., Band, S. S., & Mosavi, A. (2020). Training Multilayer

- Perceptron with Genetic Algorithms and Particle Swarm Optimization for Modeling Stock Price Index Prediction. *Entropy*, 22(11), 1–20. <https://doi.org/10.3390/e22111239>
- Emmert-Streib, F., & Dehmer, M. (2019). Evaluation of Regression Models: Model Assessment, Model Selection and Generalization Error. *Machine Learning and Knowledge Extraction*, 1(1). <https://doi.org/10.3390/make1010032>
- Erzin, Y., & Turkoz, D. (2016). Use of neural networks for the prediction of the CBR value of some Aegean sands. *Neural Computing and Applications*, 27(5), 1415–1426. <https://doi.org/10.1007/s00521-015-1943-7>
- Escamilla-García, A., Soto-Zarazúa, G. M., Toledano-Ayala, M., Rivas-Araiza, E., & Gastélum-Barrios, A. (2020). Applications of Artificial Neural Networks in Greenhouse Technology and Overview for Smart Agriculture Development. *Applied Sciences*, 10(11), Article 3835. <https://doi.org/10.3390/app10113835>
- Fikret Kurnaz, T., & Kaya, Y. (2019). Prediction of the California bearing ratio (CBR) of compacted soils by using GMDH-type neural network. *The European Physical Journal Plus*, 134(7). <https://doi.org/10.1140/epjp/i2019-12692-0>
- François Chollet. (2017). *Deep Learning with Python*. Manning Publications.
- Ghorbani, A., & Hasanzadehshooili, H. (2018). Prediction of UCS and CBR of microsilica-lime stabilized sulfate silty sand using ANN and EPR models; application to the deep soil mixing. *Soils and Foundations*, 58(1), 34–49. <https://doi.org/10.1016/j.sandf.2017.11.002>
- Godfrey, L. B., & Gashler, M. S. (2015). A continuum among logarithmic, linear, and exponential functions, and its potential to improve generalization in neural networks. *IC3K 2015 - Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, 1, 481–486. <https://doi.org/10.5220/0005635804810486>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*.
- Hernández-Sampieri, R., & Mendoza Torres, C. P. (2018). *Metodología de la Investigación: las rutas cuantitativas, cualitativa y mixta* (1st ed.). McGraw Hill.
- Hurtado de Barrera, J. (2015). *El proyecto de investigación, comprensión holística de la metodología y la investigación* (8va ed.). <https://docer.com.ar/doc/e5ne8e>

- Instituto Nacional de Estadística e Informática. (2020). *Informe Técnico N° 03 - Indicador de la Actividad Productiva Departamental: Segundo Trimestre 2020*. <https://www.inei.gov.pe/biblioteca-virtual/boletines/produccion-nacional-departamenal-9836/1/>
- Jassam Mohammed, A., Hameed Arif, M., & Adil Ali, A. (2020). A multilayer perceptron artificial neural network approach for improving the accuracy of intrusion detection systems. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 9(4), 609–615. <https://doi.org/10.11591/ijai.v9.i4.pp609-615>
- Ji, X., Li, J., Cui, Z., Li, S., Xiong, Y., Hu, J., Jiang, Y., & Zhang, C. (2020). A Prediction Method for the California Bearing Ratio of Soil-Rock Mixture Based on the Discrete Element Method and CT Scanning. *Advances in Civil Engineering*, 2020, 1–12. <https://doi.org/10.1155/2020/9794756>
- Katte, V. Y., Mfoyet, S. M., Manefouet, B., Wouatong, A. S. L., & Bezeng, L. A. (2019). eliminar. *Geotechnical and Geological Engineering*, 37(1), 217–234. <https://doi.org/10.1007/s10706-018-0604-x>
- Khalid, U., & Rehman, Z. ur. (2018). Evaluation of compaction parameters of fine-grained soils using standard and modified efforts. *International Journal of Geo-Engineering*, 9(1). <https://doi.org/10.1186/s40703-018-0083-1>
- Khuntia, S., Mujtaba, H., Patra, C., Farooq, K., Sivakugan, N., & Das, B. M. (2015). Prediction of compaction parameters of coarse grained soil using multivariate adaptive regression splines (MARS). *International Journal of Geotechnical Engineering*, 9(1), 79–88. <https://doi.org/10.1179/1939787914Y.0000000061>
- Kim, M. (2019). Supervised learning-based DDoS attacks detection: Tuning hyperparameters. *ETRI Journal*, 41(5), 560–573. <https://doi.org/10.4218/etrij.2019-0156>
- Koo, J., Zhang, J., & Chaterji, S. (2018). Tiresias: Context-sensitive approach to decipher the presence and strength of microRNA regulatory interactions. *Theranostics*, 8(1). <https://doi.org/10.7150/thno.22065>
- Lathuiliere, S., Mesejo, P., Alameda-Pineda, X., & Horaud, R. (2020). A Comprehensive Analysis of Deep Regression. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(9), 2065–2081.

- <https://doi.org/10.1109/TPAMI.2019.2910523>
- Li, T., Kong, L., & Liu, B. (2020). The California Bearing Ratio and Pore Structure Characteristics of Weakly Expansive Soil in Frozen Areas. *Applied Sciences*, 10(21), 1–22. <https://doi.org/10.3390/app10217576>
- MADe at Dutch Wikipedia. (2006). *Atterbergse grenzen*. Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Atterbergse_grenzen.PNG
- Manrique Dávila, M. (2019). *Predicción de propiedades mecánicas de suelos utilizando redes neuronales artificiales* [Tesis de grado, Universidad Continental]. <https://hdl.handle.net/20.500.12394/7191>
- Martinez Mediano, C. (2014). *Técnicas e instrumentos de recogida y analisis de datos*. UNED - Universidad Nacional de Educacion a Distancia. <https://elibro.net/es/lc/bibsipan/titulos/48726>
- Meseguer Gonzalez, P., & Badia, R. L. de M. (2017). *Inteligencia artificial*. Editorial CSIC Consejo Superior de Investigaciones Cientificas. <https://elibro.net/es/lc/bibsipan/titulos/42319>
- Mesejo, P., Ibáñez, O., Fernández-Blanco, E., Cedrón, F., Pazos, A., & Porto-Pazos, A. B. (2015). Artificial Neuron–Glia Networks Learning Approach Based on Cooperative Coevolution. *International Journal of Neural Systems*, 25(04). <https://doi.org/10.1142/S0129065715500124>
- Mina, E., Kusuma, R. I., & Ulfah, N. (2019). Utilization of steel slag and fly ash in soil stabilization and their effect to california bearing ratio (CBR) value. (Case study: Kp. Kadusentar road Medong village Mekarjaya Subdistrict Pandeglang District). *IOP Conference Series: Materials Science and Engineering*, 673(1). <https://doi.org/10.1088/1757-899X/673/1/012034>
- Ministerio de Transportes y Comunicaciones. (2014). *Manual de carreteras - Suelos, Geología, Geotécnia y Pavimentos. Sección: Suelos y Pavimentos*.
- Ministerio de Transportes y Comunicaciones. (2016). *Manual de ensayo de materiales*. https://portal.mtc.gob.pe/transportes/caminos/normas_carreteras/manuales.html
- Müller, A. C., & Guido, S. (2016). *Introduction to machine learning with Python: a guide for data scientists*. O'Reilly Media, Inc.
- Ncibi, K., Sadraoui, T., Faycel, M., & Djenina, A. (2017). A Multilayer Perceptron

- Artificial Neural Networks Based a Preprocessing and Hybrid Optimization Task for Data Mining and Classification. *International Journal of Econometrics and Financial Management*, Vol. 5, 2017, Pages 12-21, 5(1).
<https://doi.org/10.12691/ijefm-5-1-3>
- Nini, R. (2019, April). Effect of Soaking Period of Clay on its California Bearing Ratio Value. *World Congress on Civil, Structural, and Environmental Engineering*.
<https://doi.org/10.11159/icgre19.162>
- Nwankpa, C., Ijomah, W., Gachagan, A., & Marshall, S. (2018). *Activation Functions: Comparison of trends in Practice and Research for Deep Learning*. <http://arxiv.org/abs/1811.03378>
- Özgan, E., Serin, S., & Vural, İ. (2015). Multi-faceted investigation and modeling of compaction parameters for road construction. *Journal of Terramechanics*, 60, 33–42. <https://doi.org/10.1016/j.jterra.2015.02.005>
- Özyankı, A. (2019). *Soil classification by using artificial neural networks* [Master's thesis, Near East University]. <http://docs.neu.edu.tr/library/6727166826.pdf>
- Pack, C., Shin, S., Choi, H.-D., Jeon, S.-I., & Kim, J. (2016). Optimized multilayer perceptron using dynamic learning rate based microwave tomography breast cancer screening. *Proceedings of the 31st Annual ACM Symposium on Applied Computing, 04-08-April, 2171–2175*.
<https://doi.org/10.1145/2851613.2851825>
- Perez, L., Perez, R., & Seca, M. V. (2020). *Metodologia de la investigacion cientifica*. Editorial Maipue. <https://elibro.net/es/lc/bibsipan/titulos/138497>
- Raad Al-Adhadh, A., Kadhem Sakban, H., & Tawfiq Naeem, Z. (2020). Effect of Method of Soil Drying On Atterberg Limits and Soil Classification. *IOP Conference Series: Materials Science and Engineering*, 739(1).
<https://doi.org/10.1088/1757-899X/739/1/012044>
- Rajakumar, C., & Reddy babu, G. (2020). Experimental study and neural network modelling of expansive sub grade stabilized with industrial waste by-products and geogrid. *Materials Today: Proceedings*.
<https://doi.org/10.1016/j.matpr.2020.06.578>
- Ramachandran, P., Zoph, B., & Le, Q. V. (2017, October 16). Searching for Activation Functions. *6th International Conference on Learning Representations, ICLR 2018*. <http://arxiv.org/abs/1710.05941>

- Rehman, Z. U., Khalid, U., Farooq, K., & Mujtaba, H. (2017). Prediction of CBR value from index properties of different soils. *Technical Journal University of Engineering and Technology Taxila, Pakistan*, 22(2).
- Roy, S., Das, N., Kundu, M., & Nasipuri, M. (2017). Handwritten isolated Bangla compound character recognition: A new benchmark using a novel deep learning approach. *Pattern Recognition Letters*, 90, 15–21. <https://doi.org/10.1016/j.patrec.2017.03.004>
- Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., & Faubert, J. (2019). Deep learning-based electroencephalography analysis: a systematic review. *Journal of Neural Engineering*, 16(5). <https://doi.org/10.1088/1741-2552/ab260c>
- rsaad77. (2018). *Test Sieves*. Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Test_Sieves.jpg
- Saffariha, M., Jahani, A., Jahani, R., & Latif, S. (2021). Prediction of hypericin content in *Hypericum perforatum* L. in different ecological habitat using artificial neural networks. *Plant Methods*, 17(1). <https://doi.org/10.1186/s13007-021-00710-z>
- Saha, S., Paul, G. C., Pradhan, B., Abdul Maulud, K. N., & Alamri, A. M. (2021). Integrating multilayer perceptron neural nets with hybrid ensemble classifiers for deforestation probability assessment in Eastern India. *Geomatics, Natural Hazards and Risk*, 12(1), 29–62. <https://doi.org/10.1080/19475705.2020.1860139>
- Saikia, P., Baruah, R. D., Singh, S. K., & Chaudhuri, P. K. (2020). Artificial Neural Networks in the domain of reservoir characterization: A review from shallow to deep models. *Computers and Geosciences*, 135. <https://doi.org/10.1016/j.cageo.2019.104357>
- Sandoval Vallejo, E. A., & Rivera Mena, W. A. (2019). Correlación del CBR con la resistencia a la compresión inconfnada. *Ciencia e Ingeniería Neogranadina*, 29(1), 135–152. <https://doi.org/10.18359/rcin.3478>
- Shanmuganathan, S. (2016). Artificial Neural Network Modelling: An Introduction. In *Studies in Computational Intelligence* (Vol. 628, pp. 1–14). Springer Verlag. https://doi.org/10.1007/978-3-319-28495-8_1
- Sharma, S., Sharma, S., & Athaiya, A. (2020). ACTIVATION FUNCTIONS IN

- NEURAL NETWORKS. *International Journal of Engineering Applied Sciences and Technology*, 04(12), 310–316.
<https://doi.org/10.33564/IJEAST.2020.v04i12.054>
- SINGH, G., & SINGH, A. K. (2016). A study on precursors leading to geomagnetic storms using artificial neural network. *Journal of Earth System Science*, 125(5), 899–908. <https://doi.org/10.1007/s12040-016-0702-1>
- Suthar, M., & Aggarwal, P. (2018). Predicting CBR Value of Stabilized Pond Ash with Lime and Lime Sludge Using ANN and MR Models. *International Journal of Geosynthetics and Ground Engineering*, 4(1), 6.
<https://doi.org/10.1007/s40891-017-0125-3>
- Szandała, T. (2021). Review and Comparison of Commonly Used Activation Functions for Deep Neural Networks. In *Studies in Computational Intelligence* (Vol. 903, pp. 203–224). https://doi.org/10.1007/978-981-15-5495-7_11
- Taleb Bahmed, I., Harichane, K., Ghrici, M., Boukhatem, B., Rebouh, R., & Gadouri, H. (2019). Prediction of geotechnical properties of clayey soils stabilised with lime using artificial neural networks (ANNs). *International Journal of Geotechnical Engineering*, 13(2), 191–203.
<https://doi.org/10.1080/19386362.2017.1329966>
- Tenpe, A. R., & Patel, A. (2020). Application of genetic expression programming and artificial neural network for prediction of CBR. *Road Materials and Pavement Design*, 21(5), 1183–1200.
<https://doi.org/10.1080/14680629.2018.1544924>
- Tizpa, P., Jamshidi Chenari, R., Karimpour Fard, M., & Lemos Machado, S. (2015). ANN prediction of some geotechnical properties of soil from their index parameters. *Arabian Journal of Geosciences*, 8(5), 2911–2920.
<https://doi.org/10.1007/s12517-014-1304-3>
- Universidad Nacional de Ingeniería. (2019). *Tarifa de ensayos de laboratorio*. [http://www.lms.uni.edu.pe/labsuelos/tarifa/tarifa general 2019.pdf](http://www.lms.uni.edu.pe/labsuelos/tarifa/tarifa%20general%202019.pdf)
- Vafoeva, O. (2020). Hydromechanical Method of Soil Compaction. *IOP Conference Series: Materials Science and Engineering*, 883(1).
<https://doi.org/10.1088/1757-899X/883/1/012061>
- Valderrama Purizaca, F. J., Chávez Barturen, D. A., Muñoz Pérez, S. P., Tuesta-Monteza, V. A., & Mejía-Cabrera, H. I. (2020). Importance of artificial neural

- networks in civil engineering: a systematic review of the literature. *ITECKNE*, 18(1). <https://doi.org/10.15332/iteckne.v18i1.2542>
- Velasco, L. C. P., Lou, D., Paolo, G., Bryan, M., & B., F. (2018). Load Forecasting using Autoregressive Integrated Moving Average and Artificial Neural Network. *International Journal of Advanced Computer Science and Applications*, 9(7), 23–29. <https://doi.org/10.14569/IJACSA.2018.090704>
- Velasco, L. C. P., P., R., Shahin, M., F., B., & C., J. (2019). Performance Analysis of Multilayer Perceptron Neural Network Models in Week-Ahead Rainfall Forecasting. *International Journal of Advanced Computer Science and Applications*, 10(3), 578–588. <https://doi.org/10.14569/IJACSA.2019.0100374>
- Velasco, L. C. P., Villezas, C. R., Palahang, P. N. C., & Dagaang, J. A. A. (2015). Next day electric load forecasting using Artificial Neural Networks. *2015 International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, 1–6. <https://doi.org/10.1109/HNICEM.2015.7393166>
- Verma, G., & Kumar, B. (2020). Prediction of compaction parameters for fine-grained and coarse-grained soils: a review. *International Journal of Geotechnical Engineering*, 14(8), 970–977. <https://doi.org/10.1080/19386362.2019.1595301>
- Vijayashanthar, V., Qiao, J., Zhu, Z., Entwistle, P., & Yu, G. (2018). Modeling Fecal Indicator Bacteria in Urban Waterways Using Artificial Neural Networks. *Journal of Environmental Engineering*, 144(6). [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0001377](https://doi.org/10.1061/(ASCE)EE.1943-7870.0001377)
- Wang, Yingying, Li, Y., Song, Y., & Rong, X. (2020). The Influence of the Activation Function in a Convolution Neural Network Model of Facial Expression Recognition. *Applied Sciences*, 10(5). <https://doi.org/10.3390/app10051897>
- Wang, Yixiang, Liu, J., Misic, J., Misic, V. B., Lv, S., & Chang, X. (2019). Assessing Optimizer Impact on DNN Model Sensitivity to Adversarial Examples. *IEEE Access*, 7, 152766–152776. <https://doi.org/10.1109/ACCESS.2019.2948658>
- Ward, R., Wu, X., & Bottou, L. (2018). AdaGrad stepsizes: Sharp convergence over nonconvex landscapes, from any initialization. *36th International Conference on Machine Learning, ICML 2019, 2019-June*, 11574–11583.

<http://arxiv.org/abs/1806.01811>

- Yi, D., Ahn, J., & Ji, S. (2020). An Effective Optimization Method for Machine Learning Based on ADAM. *Applied Sciences*, 10(3). <https://doi.org/10.3390/app10031073>
- YILDIRIM, S. (2019). Improving word embeddings projection for Turkish hypernym extraction. *TURKISH JOURNAL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES*, 27(6), 4418–4428. <https://doi.org/10.3906/elk-1903-65>
- Zhang, X., Chen, X., Wang, J., Zhan, Z., & Li, J. (2018). Verifiable privacy-preserving single-layer perceptron training scheme in cloud computing. *Soft Computing*, 22(23), 7719–7732. <https://doi.org/10.1007/s00500-018-3233-7>
- Zhou, B., & Lu, N. (2021). Correlation between Atterberg Limits and Soil Adsorptive Water. *Journal of Geotechnical and Geoenvironmental Engineering*, 147(2). [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0002463](https://doi.org/10.1061/(ASCE)GT.1943-5606.0002463)

ANEXOS

Anexo 1. Formato de la ficha de registro de datos para el conjunto de datos

Los resultados de los ensayos de laboratorio fueron registrados en la siguiente tabla:

Muestra	Contenido de humedad (%)	Contenido de sales (%)	Análisis granulométrico (% que pasa)			Límites de Atterberg			Clasificación		Compactación		CBR	
			3 in	n.º 4	n.º 200	LL	PL	PI	SUCS	AASHTO	MDD	OMC	al 95 %	al 100 %
1														
2														
:														

Se da el formato que se crea conveniente para ingresar al programa Spyder. En este caso se utilizó la siguiente tabla:

Muestra	H	S	G	A	F	LL	PL	PI	SUCS	AASHTO	MDD	OMC	CBR95	CBR100
1														
2														
:														

Anexo 2. Formato de la ficha bibliográfica para el número de entradas y el número de salidas según la literatura

La siguiente tabla se utilizó para registrar las variables de entrada (X) y salida (y) según la literatura:

	Autor	H	S	G	A	F	LL	PL	PI	O	MDD	OMC	CBR
1			X			X							y
2													
			⋮										

Se utilizó una tabla similar para registrar las variables de entrada y salida utilizadas en esta investigación para cada modelo:

Modelo	H	S	G	A	F	LL	PL	PI	MDD	OMC	CBR95	CBR100
1			X									y
2												
			⋮									

Anexo 3. Ficha de registro de datos para los hiperparámetros asociados al modelo e hiperparámetros asociados al descenso del gradiente

En la tabla se registran cada uno de los hiperparámetros probados para cada modelo. Además, en ella se registran los resultados obtenidos por cada modelo y la posición del modelo según su rendimiento.

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
1												
2												
3												
4												
5												
6												
8												
9												
10												
...												

Anexo 4. Tabla de registro de datos para el rendimiento de los modelos

ANN	MSE	MAE	R
Propiedad del suelo 1			
1			
2			
Propiedad del suelo 2			
3			
4			
Propiedad del suelo 3			
5			
Propiedad del suelo 4			
6			

Anexo 5. Características del conjunto de datos

Clasificación USCS				
Gravas	Frecuencias	Frecuencia acumulada	Porcentaje	Porcentaje acumulado
GC	7	7	20.59	20.59
GC-GM	5	12	14.71	35.30
GM	17	29	50.00	85.30
GM-GC	1	30	2.94	88.24
GP-GC	2	32	5.88	94.12
GP-GM	1	33	2.94	97.06
GW-GM	1	34	2.94	100.00
Total	34		100.00	

Clasificación USCS				
Arenas	Frecuencias	Frecuencia acumulada	Porcentaje	Porcentaje acumulado
SC	11	11	8.66	8.66
SC-SM	13	24	10.24	18.90
SM	72	96	56.69	75.59
SM-SC	11	107	8.66	84.25
SP	13	120	10.24	94.49
SP-SM	7	127	5.51	100.00
Total	127		100.00	

Clasificación USCS				
Limos y arcillas	Frecuencias	Frecuencia acumulada	Porcentaje	Porcentaje acumulado
CH	6	6	4.84	4.84
CL	92	98	74.19	79.03
CL-ML	9	107	7.26	86.29
MH	2	109	1.61	87.90
ML	11	120	8.87	96.77
ML-CL	4	124	3.23	100.00
Total	124		100.00	

Clasificación AASTHO				
Materiales granulares	Frecuencias	Frecuencia acumulada	Porcentaje	Porcentaje acumulado
A-1-a	15	15	9.55	9.55
A-1-b	13	28	8.28	17.83
A-2-4	104	132	66.24	84.07
A-2-6	8	140	5.10	89.17
A-2-7	1	141	0.64	89.81
A-3	16	157	10.19	100.00
Total	157		100.00	

Clasificación AASTHO				
Material limo arcilloso	Frecuencias	Frecuencia acumulada	Porcentaje	Porcentaje acumulado
A-4	39	39	30.47	30.47
A-6	67	106	52.34	82.81
A-7-6	22	128	17.19	100.00
Total	128		100.00	

Rango del conjunto de datos			
Ensayo	Mínimo	Máximo	
Contenido de humedad (%)	1.10	36.07	
Contenido de sales (%)	0.00	2.52	
Contenido de grava (%)	0.00	81.90	
Contenido de arena (%)	1.97	100.00	
Contenido de finos (%)	0.00	97.69	
Límite líquido (%)	0.00	59.58	
Límite plástico (%)	0.00	27.93	
Índice de plasticidad (%)	0.00	37.44	
MDD (g/cm ³)	1.56	2.21	
Óptimo contenido de humedad (%)	5.85	22.33	
CBR al 95 % de la máxima densidad seca (%)	2.81	72.00	
CBR al 100 % de la máxima densidad seca (%)	4.20	86.00	

Anexo 6. Correlación entre las variables de entradas y las variables de salida

		1	2	3	4	5	6	7	8	9	10	11	12
1. Contenido de humedad	rs	1.000											
	Sig.												
2. Contenido de sales	rs	0.249*	1.000										
	Sig.	0.000											
3. Contenido de grava	rs	-0.216*	-0.220*	1.000									
	Sig.	0.000	0.000										
4. Contenido de arena	rs	-0.428*	-0.084	-0.265*	1.000								
	Sig.	0.000	0.155	0.000									
5. Contenido de finos	rs	0.429*	0.158*	-0.118	-0.816*	1.000							
	Sig.	0.000	0.008	0.047	0.000								
6. Límite líquido	rs	0.550*	0.145	-0.069	-0.717*	0.693*	1.000						
	Sig.	0.000	0.014	0.244	0.000	0.000							

Nota. rs = coeficiente de correlación de Spearman; Sig. = nivel de significancia.

* $p < 0.01$, la correlación es significativa en el nivel 0.01 (2 colas).

7. Límite plástico	rs	0.509*	0.021	-0.056	-0.532*	0.450*	0.839*	1.000					
	Sig.	0.000	0.718	0.350	0.000	0.000	0.000						
8. Índice de plasticidad	rs	0.557*	0.229*	-0.041	-0.734*	0.736*	0.931*	0.693*	1.000				
	Sig.	0.000	0.000	0.490	0.000	0.000	0.000	0.000					
9. Máxima densidad seca	rs	-0.386*	-0.214*	0.572*	0.101	-0.336*	-0.369*	-0.359*	-0.375*	1.000			
	Sig.	0.000	0.000	0.000	0.089	0.000	0.000	0.000	0.000				
10. Óptimo contenido de humedad	rs	0.506*	0.063	-0.373*	-0.369*	0.500*	0.613*	0.523*	0.606*	-0.715*	1.000		
	Sig.	0.000	0.292	0.000	0.000	0.000	0.000	0.000	0.000	0.000			
11. CBR al 95 %	rs	-0.186*	-0.168*	0.391*	0.213*	-0.570*	-0.208*	-0.017	-0.291*	0.371*	-0.335*	1.000	
	Sig.	0.002	0.004	0.000	0.000	0.000	0.000	0.769	0.000	0.000	0.000		
12. CBR al 100 %	rs	-0.217*	-0.222*	0.401*	0.254*	-0.609*	-0.238*	-0.057	-0.318*	0.381*	-0.343*	0.947*	1.000
	Sig.	0.000	0.000	0.000	0.000	0.000	0.000	0.338	0.000	0.000	0.000	0.000	

Nota. rs = coeficiente de correlación de Spearman; Sig. = nivel de significancia.

* $p < 0.01$, la correlación es significativa en el nivel 0.01 (2 colas).

Anexo 7. Modelos entrenados para predecir el CBR al 100 % de la máxima densidad seca utilizando 8 variables de entrada

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
1	4	3	17	8E-05	0.3	32	1414	1.14E-02	8.51E-03	8.16E-03	9.37E-03	31
2	2	5	18	1E-04	0.2	16	820	5.62E-03	5.58E-03	5.16E-03	5.47E-03	1
3	4	3	1	1E-01	0.4	16	999	2.50E-02	2.65E-02	2.67E-02	2.61E-02	73
4	3	10	18	2E-03	0	32	999	1.31E-02	1.35E-02	8.66E-03	1.17E-02	37
5	3	19	5	3E-06	0.5	128	653	6.78E-02	4.56E-02	4.87E-02	5.41E-02	90
6	5	9	1	7E-01	0.2	64	487	2.74E-02	2.95E-02	2.71E-02	2.80E-02	77
7	4	18	6	1E-02	0.4	64	778	1.83E-02	1.52E-02	1.90E-02	1.76E-02	51
8	6	18	9	1E-02	0.3	128	977	1.77E-02	1.71E-02	1.71E-02	1.73E-02	49
9	3	15	17	2E-03	0.1	128	595	9.35E-03	6.76E-03	6.56E-03	7.53E-03	19
10	2	10	18	6E-04	0.2	64	802	5.95E-03	9.11E-03	5.80E-03	6.99E-03	13
11	5	20	3	2E-04	0.1	16	100	1.42E-02	1.73E-02	1.53E-02	1.56E-02	45
12	4	20	20	1E+00	0.5	128	100	1.48E+00	1.54E+00	1.11E+00	1.38E+00	99
13	2	1	2	1E-06	0.5	16	1451	8.09E-02	7.73E-02	1.68E-01	1.09E-01	95

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
14	2	17	12	3E-06	0.3	64	104	1.20E-01	9.62E-02	1.04E-01	1.07E-01	94
15	5	18	7	1E-02	0.3	64	795	1.65E-02	1.65E-02	1.58E-02	1.63E-02	47
16	2	1	1	2E-05	0	128	1500	1.40E-01	1.25E-01	1.05E-01	1.22E-01	96
17	6	1	9	1E-03	0.5	16	1500	2.01E-02	2.13E-02	2.05E-02	2.06E-02	58
18	6	20	11	3E-03	0	64	1500	7.64E-03	8.58E-03	8.73E-03	8.30E-03	25
19	6	20	9	1E-04	0.5	128	100	3.67E-02	3.43E-02	3.70E-02	3.61E-02	82
20	6	20	20	1E-06	0.5	128	1500	4.95E-02	4.63E-02	4.63E-02	4.73E-02	87
21	6	20	20	7E-05	0.5	32	1500	1.37E-02	1.81E-02	1.66E-02	1.61E-02	46
22	6	20	2	1E-03	0.5	128	100	2.35E-02	2.36E-02	2.35E-02	2.35E-02	66
23	6	1	16	1E-03	0	16	1500	1.52E-02	1.05E-02	9.70E-03	1.18E-02	38
24	2	1	20	1E-05	0	16	100	5.18E-02	5.68E-02	4.70E-02	5.19E-02	88
25	2	20	9	5E-03	0.5	128	100	7.60E-03	7.15E-03	7.85E-03	7.54E-03	20
26	2	20	5	5E-04	0	128	1500	7.39E-03	5.67E-03	5.40E-03	6.17E-03	8
27	6	20	18	2E-04	0	128	100	1.12E-02	1.08E-02	1.15E-02	1.12E-02	36
28	2	20	14	3E-03	0.5	16	1500	6.71E-03	2.37E-02	8.13E-03	1.29E-02	39

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
29	2	1	20	3E-04	0.4	128	1500	2.25E-02	8.43E-03	1.16E-02	1.42E-02	43
30	2	20	1	1E-02	0	128	1500	1.75E-02	1.55E-02	1.91E-02	1.74E-02	50
31	2	20	15	3E-04	0	128	1500	5.88E-03	6.82E-03	5.28E-03	6.00E-03	7
32	2	20	1	4E-04	0.4	128	1500	1.88E-02	2.02E-02	2.10E-02	2.00E-02	57
33	2	1	5	4E-03	0	64	1500	1.64E-02	2.51E-02	1.22E-02	1.80E-02	56
34	2	1	12	4E-03	0	128	100	1.78E-02	1.34E-02	1.89E-02	1.67E-02	48
35	2	20	17	2E-04	0.5	128	1500	5.67E-03	6.06E-03	5.83E-03	5.81E-03	5
36	2	20	14	1E-03	0	32	100	7.29E-03	6.26E-03	5.23E-03	6.27E-03	9
37	2	20	8	1E-06	0	128	1500	7.33E-02	1.09E-01	8.57E-02	8.93E-02	93
38	2	20	1	3E-01	0	128	1500	2.46E-02	2.32E-02	2.37E-02	2.38E-02	69
39	2	20	18	1E-04	0	16	1500	4.82E-03	7.31E-03	7.34E-03	6.50E-03	11
40	2	20	15	1E-04	0.5	16	100	9.15E-03	9.58E-03	1.03E-02	9.71E-03	33
41	6	20	1	4E-02	0	128	100	2.36E-02	2.35E-02	2.34E-02	2.35E-02	65
42	2	20	9	4E-04	0	16	1398	7.81E-03	7.22E-03	1.50E-02	1.00E-02	34
43	2	1	19	8E-05	0.5	128	100	5.02E-02	3.94E-02	6.77E-02	5.23E-02	89

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
44	6	20	17	2E-04	0.1	16	1500	7.45E-03	8.37E-03	1.16E-02	9.10E-03	29
45	2	20	18	6E-04	0	16	1500	6.98E-03	6.41E-03	7.08E-03	6.83E-03	12
46	6	20	20	6E-04	0	16	1500	7.85E-03	6.98E-03	1.16E-02	8.80E-03	27
47	2	20	8	5E-03	0	16	1500	1.03E-01	9.39E-03	1.07E-02	4.13E-02	83
48	6	20	14	1E-03	0.5	128	1500	1.72E-02	1.75E-02	1.86E-02	1.77E-02	53
49	6	20	17	3E-04	0	16	1500	5.48E-03	8.45E-03	7.78E-03	7.24E-03	14
50	6	2	10	4E-03	0.5	128	100	2.29E-02	2.32E-02	2.34E-02	2.32E-02	62
51	2	20	5	3E-04	0.5	16	1500	8.91E-03	1.09E-02	8.91E-03	9.59E-03	32
52	2	20	4	3E-02	0.5	128	100	1.74E-02	1.32E-02	1.01E-02	1.35E-02	42
53	2	20	20	3E-04	0	16	1500	6.52E-03	1.13E-02	5.80E-03	7.90E-03	23
54	2	20	15	1E-04	0	16	1500	1.13E-02	7.33E-03	6.17E-03	8.30E-03	26
55	2	20	6	6E-04	0	128	100	9.26E-03	8.18E-03	1.06E-02	9.35E-03	30
56	2	20	17	7E-04	0.5	16	1500	7.83E-03	2.42E-02	7.32E-03	1.32E-02	41
57	3	15	18	9E-01	0.2	64	1213	2.39E-02	2.40E-02	2.34E-02	2.38E-02	68
58	2	20	20	8E-04	0.4	128	1500	7.67E-03	6.36E-03	8.19E-03	7.32E-03	17

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
59	4	8	1	5E-01	0	64	560	2.70E-02	2.55E-02	2.60E-02	2.61E-02	74
60	3	12	20	2E-01	0	128	914	2.05E-02	2.31E-02	2.08E-02	2.15E-02	59
61	6	1	1	3E-02	0	128	1500	2.36E-02	2.37E-02	2.36E-02	2.36E-02	67
62	2	17	15	1E-03	0.2	32	691	9.05E-03	6.33E-03	6.41E-03	7.29E-03	16
63	2	15	20	1E-02	0.5	64	776	7.05E-03	8.88E-03	1.50E-02	1.03E-02	35
64	2	20	20	4E-02	0	128	963	1.19E-02	2.87E-02	1.29E-02	1.79E-02	54
65	5	18	8	4E-05	0	32	576	5.49E-03	6.05E-03	5.67E-03	5.76E-03	3
66	3	9	19	2E-03	0	16	1001	6.71E-03	9.13E-03	2.30E-02	1.30E-02	40
67	5	20	15	6E-06	0.3	128	748	3.24E-02	3.08E-02	3.60E-02	3.30E-02	80
68	2	20	1	2E-02	0	128	110	1.75E-02	2.77E-02	2.10E-02	2.21E-02	60
69	3	17	1	1E-06	0.5	64	266	4.89E-02	9.60E-02	4.81E-02	6.38E-02	92
70	2	20	1	2E-06	0	32	100	5.55E-02	7.54E-02	3.18E-01	1.51E-01	97
71	2	20	20	4E-02	0	16	408	5.26E-02	1.85E-02	8.69E-03	2.65E-02	75
72	6	1	20	1E+00	0.3	64	880	4.40E-01	3.00E-02	2.35E-02	1.66E-01	98
73	6	20	7	3E-04	0	64	1500	7.01E-03	5.73E-03	1.18E-02	8.22E-03	24

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
74	2	1	20	1E+00	0	128	1413	2.43E-02	2.34E-02	2.56E-02	2.44E-02	70
75	6	20	20	1E-06	0.5	64	1014	4.37E-02	5.04E-02	4.43E-02	4.61E-02	85
76	2	20	18	1E-02	0	32	100	8.24E-03	6.30E-03	8.75E-03	7.79E-03	22
77	5	19	14	3E-02	0.3	32	500	2.27E-02	2.38E-02	2.01E-02	2.22E-02	61
78	2	20	20	9E-06	0.5	16	359	1.88E-02	1.76E-02	1.75E-02	1.79E-02	55
79	2	20	20	1E-04	0.5	16	596	6.12E-03	5.36E-03	5.82E-03	5.77E-03	4
80	2	1	20	3E-03	0.5	16	1500	1.29E-02	1.57E-02	1.79E-02	1.54E-02	44
81	6	20	20	1E-04	0	128	1088	5.28E-03	9.09E-03	5.08E-03	6.47E-03	10
82	5	20	7	1E-06	0.4	64	991	6.48E-02	6.65E-02	4.51E-02	5.89E-02	91
83	2	20	20	4E-05	0.2	16	689	5.60E-03	5.50E-03	5.58E-03	5.56E-03	2
84	5	1	1	1E-06	0	32	1498	5.11E-02	4.44E-02	4.59E-02	4.71E-02	86
85	4	20	16	1E+00	0.3	32	1069	2.50E-02	2.42E-02	2.43E-02	2.45E-02	71
86	4	1	20	1E-06	0.2	32	813	3.47E-02	4.11E-02	5.98E-02	4.52E-02	84
87	4	1	1	4E-03	0.5	32	1466	2.36E-02	2.35E-02	2.36E-02	2.35E-02	64
88	4	20	20	1E-06	0.3	16	940	3.01E-02	3.02E-02	3.45E-02	3.16E-02	78

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
89	6	20	20	2E-01	0.5	128	495	2.34E-02	2.35E-02	2.34E-02	2.34E-02	63
90	5	1	1	1E+00	0.1	128	1489	2.73E-02	2.88E-02	2.65E-02	2.75E-02	76
91	4	20	20	5E-04	0	32	1500	7.55E-03	1.01E-02	9.01E-03	8.88E-03	28
92	6	20	20	1E-04	0	128	1346	6.30E-03	5.99E-03	5.50E-03	5.93E-03	6
93	6	20	20	2E-04	0	16	937	6.03E-03	9.25E-03	6.89E-03	7.38E-03	18
94	6	2	19	1E+00	0.2	32	208	4.81E+00	2.70E+00	1.83E+00	3.14E+00	100
95	2	20	15	3E-02	0	16	427	1.93E-02	7.46E-03	7.03E-02	3.24E-02	79
96	2	20	20	4E-04	0	128	127	6.85E-03	7.05E-03	7.78E-03	7.25E-03	15
97	6	1	20	1E-06	0.3	32	1206	3.52E-02	3.16E-02	3.46E-02	3.39E-02	81
98	2	1	20	1E-02	0.5	128	895	2.22E-02	1.45E-02	1.61E-02	1.76E-02	52
99	6	1	20	4E-06	0.1	32	1111	2.46E-02	2.42E-02	2.61E-02	2.49E-02	72
100	6	20	20	2E-04	0	32	1450	6.17E-03	6.43E-03	1.00E-02	7.57E-03	21

Anexo 8. Modelos entrenados para predecir el CBR al 95 % de la máxima densidad seca utilizando 8 variables de entrada

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
1	4	7	13	3E-06	0.4	32	1342	1.94E-02	1.66E-02	1.96E-02	1.86E-02	89
2	5	6	15	1E-06	0	64	138	3.38E-02	4.56E-02	2.71E-02	3.53E-02	95
3	3	19	3	2E-05	0.3	64	528	1.96E-02	2.29E-02	2.00E-02	2.08E-02	91
4	4	4	12	1E-06	0.4	16	1289	2.01E-02	2.10E-02	2.19E-02	2.10E-02	92
5	3	14	17	1E-06	0.5	32	538	2.81E-02	3.76E-02	4.35E-02	3.64E-02	96
6	3	8	9	5E-01	0	16	1188	2.39E-02	1.92E-02	1.84E-02	2.05E-02	90
7	3	6	8	4E-05	0.2	64	780	1.13E-02	1.24E-02	9.84E-03	1.12E-02	58
8	6	15	5	6E-06	0.4	128	128	3.10E-02	5.23E-02	3.14E-02	3.82E-02	97
9	5	4	11	6E-02	0	32	168	1.19E-02	1.20E-02	1.55E-02	1.31E-02	64
10	6	8	13	2E-01	0.4	32	1017	1.59E-02	1.60E-02	1.58E-02	1.60E-02	82
11	3	16	6	1E-03	0	32	997	7.42E-03	7.54E-03	7.62E-03	7.57E-03	41
12	3	12	8	1E-03	0.5	64	847	8.50E-03	8.50E-03	7.56E-03	8.22E-03	44
13	5	13	7	2E-04	0.4	64	1500	1.16E-02	1.28E-02	1.21E-02	1.21E-02	63

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
14	5	4	12	8E-03	0.4	64	1500	1.08E-02	1.23E-02	1.12E-02	1.13E-02	59
15	6	9	3	8E-03	0	128	717	1.04E-02	1.11E-02	1.16E-02	1.10E-02	57
16	4	19	16	2E-04	0.2	32	788	6.79E-03	5.15E-03	4.92E-03	5.61E-03	22
17	5	17	8	5E-04	0.4	16	100	1.22E-02	1.21E-02	1.19E-02	1.21E-02	62
18	3	2	18	1E+00	0.3	32	100	1.55E-02	1.56E-02	1.55E-02	1.56E-02	77
19	5	3	17	1E-03	0	16	100	5.75E-03	5.41E-03	6.82E-03	6.00E-03	26
20	2	1	20	5E-04	0	16	584	6.64E-03	6.11E-03	1.03E-02	7.69E-03	42
21	5	20	14	2E-05	0	64	1500	5.47E-03	4.67E-03	5.27E-03	5.15E-03	13
22	6	20	20	1E-04	0	32	1500	6.08E-03	5.34E-03	5.42E-03	5.60E-03	21
23	5	20	4	2E-02	0.5	16	117	1.57E-02	1.53E-02	1.55E-02	1.56E-02	76
24	4	9	4	1E+00	0.5	32	1456	2.50E-02	2.06E-02	2.15E-02	2.24E-02	93
25	3	10	7	1E-06	0	64	1500	4.25E-02	4.78E-02	4.32E-02	4.45E-02	98
26	4	6	20	2E-04	0.5	16	1285	7.51E-03	8.75E-03	8.06E-03	8.14E-03	43
27	5	18	20	2E-05	0	16	1500	4.56E-03	4.46E-03	4.08E-03	4.35E-03	3
28	6	20	20	3E-03	0	128	100	4.48E-03	5.21E-03	7.84E-03	5.84E-03	24

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
29	4	1	20	1E-05	0.1	16	1500	1.28E-02	8.45E-03	9.81E-03	1.04E-02	53
30	4	6	20	2E-03	0.2	16	580	5.69E-03	6.76E-03	6.00E-03	6.16E-03	27
31	3	9	20	5E-05	0	16	1144	3.89E-03	5.02E-03	4.50E-03	4.49E-03	5
32	2	20	20	7E-05	0	32	100	9.51E-03	9.52E-03	8.90E-03	9.34E-03	51
33	4	20	1	8E-04	0.2	32	100	1.54E-02	1.54E-02	1.52E-02	1.53E-02	72
34	5	20	20	1E-02	0.2	128	1500	6.14E-03	6.71E-03	9.55E-03	7.50E-03	39
35	2	20	20	2E-03	0.1	16	1500	6.13E-03	1.87E-02	6.80E-03	1.05E-02	56
36	5	1	20	4E-03	0.3	128	100	1.21E-02	9.66E-03	9.75E-03	1.05E-02	54
37	5	20	20	2E-01	0	128	100	1.55E-02	1.53E-02	1.28E-02	1.46E-02	70
38	2	11	20	3E-02	0.5	128	100	5.07E-03	8.15E-03	5.57E-03	6.28E-03	28
39	5	13	19	4E-05	0.2	128	1457	6.34E-03	6.62E-03	6.72E-03	6.54E-03	30
40	2	1	1	2E-03	0	16	1034	1.15E-02	1.77E-02	1.41E-02	1.44E-02	67
41	3	1	20	1E-04	0.2	128	100	1.94E-02	1.72E-02	1.57E-02	1.75E-02	87
42	4	16	10	3E-05	0	16	1500	4.87E-03	5.63E-03	5.00E-03	5.18E-03	14
43	3	20	20	4E-05	0.1	32	1333	4.22E-03	3.93E-03	4.22E-03	4.11E-03	1

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
44	4	20	1	5E-03	0.2	128	1500	1.55E-02	1.48E-02	1.37E-02	1.46E-02	71
45	2	14	20	2E-02	0	16	100	8.24E-03	4.03E-03	8.91E-03	7.05E-03	38
46	2	13	16	5E-03	0	128	100	5.10E-03	4.54E-03	5.13E-03	4.94E-03	11
47	2	16	20	1E-03	0	128	1091	6.33E-03	5.15E-03	6.42E-03	5.98E-03	25
48	2	1	20	1E+00	0	128	1500	1.74E-02	1.88E-02	1.57E-02	1.73E-02	86
49	3	20	20	1E+00	0.5	16	1500	2.16E-02	2.60E-02	1.98E-02	2.25E-02	94
50	6	19	20	2E-03	0.5	64	1452	1.12E-02	1.13E-02	1.25E-02	1.16E-02	61
51	6	13	20	4E-04	0	16	440	6.31E-03	4.62E-03	6.48E-03	5.80E-03	23
52	5	1	20	1E-01	0.5	32	100	1.60E-02	1.61E-02	1.68E-02	1.63E-02	83
53	4	19	3	8E-01	0.2	128	130	1.58E-02	1.59E-02	1.57E-02	1.58E-02	80
54	5	1	1	1E+00	0.4	128	100	1.56E-02	1.54E-02	1.59E-02	1.56E-02	78
55	2	2	17	3E-05	0.2	16	157	1.79E-02	1.66E-02	2.12E-02	1.86E-02	88
56	5	1	20	2E-05	0.5	16	1500	1.42E-02	1.48E-02	1.41E-02	1.44E-02	68
57	2	17	19	2E-03	0.5	32	104	4.61E-03	4.64E-03	4.46E-03	4.57E-03	6
58	4	10	20	3E-02	0	64	1500	7.04E-03	7.39E-03	5.80E-03	6.72E-03	36

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
59	4	13	20	6E-03	0.5	32	445	8.06E-03	8.19E-03	8.42E-03	8.23E-03	45
60	4	20	20	5E-02	0.5	64	1500	1.57E-02	1.61E-02	1.57E-02	1.59E-02	81
61	3	1	5	9E-05	0.1	16	1410	9.59E-03	7.61E-03	8.13E-03	8.42E-03	47
62	4	1	20	4E-02	0.5	64	1500	1.54E-02	1.55E-02	1.54E-02	1.55E-02	74
63	6	20	20	1E+00	0	128	1500	5.95E+00	4.41E+00	6.86E+00	5.75E+00	100
64	4	20	3	8E-04	0.3	64	739	1.03E-02	9.82E-03	1.11E-02	1.04E-02	52
65	6	1	1	5E-02	0.5	128	825	1.57E-02	1.53E-02	1.55E-02	1.54E-02	73
66	2	9	18	5E-03	0.3	128	363	6.50E-03	8.73E-03	4.56E-03	6.58E-03	31
67	2	20	20	3E-03	0.3	128	100	4.40E-03	4.70E-03	4.14E-03	4.37E-03	4
68	2	20	11	8E-03	0.3	32	551	6.45E-03	5.98E-03	7.54E-03	6.66E-03	34
69	2	14	15	4E-03	0.4	128	100	4.80E-03	5.04E-03	4.24E-03	4.72E-03	8
70	2	20	15	5E-03	0.3	16	1115	1.33E-02	5.28E-03	6.92E-03	8.52E-03	48
71	2	20	20	4E-03	0.5	128	100	4.21E-03	5.19E-03	4.47E-03	4.64E-03	7
72	2	16	2	7E-03	0.1	32	100	7.91E-03	9.69E-03	7.57E-03	8.32E-03	46
73	2	9	16	9E-03	0.3	32	100	5.68E-03	5.19E-03	3.91E-03	4.94E-03	10

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
74	4	14	11	5E-03	0.2	128	109	6.04E-03	5.12E-03	4.91E-03	5.37E-03	18
75	4	9	18	5E-03	0.5	128	100	8.93E-03	9.21E-03	9.16E-03	9.13E-03	50
76	2	15	11	1E-02	0.5	128	1297	5.58E-03	8.62E-03	5.94E-03	6.70E-03	35
77	2	20	9	9E-03	0.3	16	1394	1.21E-02	6.94E-03	6.69E-03	8.56E-03	49
78	2	20	20	1E-02	0.5	128	100	5.45E-03	5.37E-03	4.15E-03	4.98E-03	12
79	2	12	20	1E-02	0.3	128	100	4.68E-03	4.73E-03	6.63E-03	5.35E-03	17
80	2	19	13	9E-03	0.1	32	100	5.92E-03	5.34E-03	4.69E-03	5.30E-03	15
81	2	19	8	2E-02	0.1	32	100	1.05E-02	7.51E-03	1.35E-02	1.05E-02	55
82	2	4	12	1E-06	0.4	64	462	1.21E-01	3.32E-02	9.66E-02	8.24E-02	99
83	2	10	18	3E-02	0.5	32	959	9.95E-03	5.67E-03	5.27E-03	6.95E-03	37
84	2	2	19	1E-02	0.4	32	100	1.03E-02	5.56E-03	6.92E-03	7.57E-03	40
85	3	1	20	3E-02	0.5	128	100	1.46E-02	1.22E-02	1.52E-02	1.40E-02	65
86	6	14	20	9E-02	0.5	16	1329	1.65E-02	1.72E-02	1.58E-02	1.65E-02	84
87	4	17	8	1E-02	0.4	64	1500	1.11E-02	1.24E-02	1.08E-02	1.15E-02	60
88	5	14	16	3E-04	0	128	130	6.41E-03	4.92E-03	5.00E-03	5.46E-03	19

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
89	6	4	4	2E-04	0.3	64	841	1.38E-02	1.48E-02	1.39E-02	1.41E-02	66
90	6	14	20	9E-05	0.2	32	1329	4.70E-03	5.51E-03	5.76E-03	5.33E-03	16
91	5	17	17	6E-04	0	128	393	6.30E-03	5.74E-03	4.44E-03	5.49E-03	20
92	5	20	20	3E-04	0	32	1500	5.91E-03	7.62E-03	6.22E-03	6.59E-03	32
93	4	14	20	7E-03	0.2	16	184	6.55E-03	6.81E-03	6.11E-03	6.50E-03	29
94	4	11	20	2E-04	0.2	64	777	5.24E-03	5.00E-03	4.39E-03	4.83E-03	9
95	5	7	6	4E-04	0.5	16	540	1.41E-02	1.48E-02	1.47E-02	1.44E-02	69
96	6	16	4	2E-04	0	64	954	4.72E-03	5.83E-03	9.14E-03	6.60E-03	33
97	6	1	1	4E-02	0.5	32	100	1.58E-02	1.52E-02	1.56E-02	1.55E-02	75
98	6	8	13	1E-01	0.4	32	1132	1.68E-02	1.76E-02	1.63E-02	1.69E-02	85
99	5	20	20	2E-04	0	16	100	4.45E-03	4.30E-03	4.24E-03	4.31E-03	2
100	4	20	1	2E-01	0.2	128	100	1.60E-02	1.54E-02	1.59E-02	1.57E-02	79

Anexo 9. Modelos entrenados para predecir el CBR al 100 % de la máxima densidad seca utilizando 10 variables de entrada

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
1	2	14	18	3E-06	0.1	32	715	5.61E-02	8.73E-02	6.65E-02	6.94E-02	92
2	4	2	7	4E-05	0.4	32	423	2.30E-02	2.46E-02	2.37E-02	2.38E-02	64
3	4	8	14	6E-03	0.4	16	1165	1.07E-02	1.11E-02	8.43E-03	1.01E-02	26
4	4	2	9	1E-02	0.2	32	431	1.09E-02	1.46E-02	9.31E-03	1.16E-02	34
5	2	16	4	1E-06	0.5	32	963	1.11E-01	1.27E-01	5.93E-02	9.89E-02	93
6	6	1	4	5E-01	0.4	64	786	2.87E-02	2.82E-02	2.72E-02	2.80E-02	75
7	2	19	11	3E-01	0.5	64	1416	2.34E-02	2.71E-02	2.71E-02	2.58E-02	71
8	2	6	16	1E-06	0.1	32	1413	5.51E-02	4.21E-02	7.35E-02	5.70E-02	91
9	4	15	7	1E-03	0.3	32	1400	9.74E-03	8.81E-03	8.52E-03	9.01E-03	23
10	1	1	15	5E-03	0.2	128	623	1.59E-02	1.80E-02	1.68E-02	1.68E-02	48
11	6	16	14	6E-03	0	16	216	6.00E-03	6.12E-03	4.92E-03	5.70E-03	10
12	6	16	1	8E-03	0	16	551	2.34E-02	2.37E-02	2.19E-02	2.30E-02	59
13	5	1	1	9E-04	0.4	32	1188	2.33E-02	2.34E-02	2.34E-02	2.34E-02	62

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
14	6	20	18	4E-04	0.3	64	1099	1.13E-02	1.00E-02	1.16E-02	1.10E-02	33
15	6	20	20	3E-03	0.5	32	1500	1.79E-02	1.76E-02	1.93E-02	1.83E-02	51
16	6	20	15	2E-02	0	32	100	5.22E-03	7.55E-03	6.67E-03	6.49E-03	12
17	5	18	20	3E-03	0	64	100	5.63E-03	4.70E-03	4.29E-03	4.88E-03	4
18	6	11	20	3E-02	0	32	100	5.51E-03	5.50E-03	5.05E-03	5.35E-03	7
19	2	1	20	1E+00	0	16	1436	4.11E-02	2.69E-02	3.09E-02	3.29E-02	84
20	6	12	18	6E-04	0.1	16	100	6.64E-03	5.33E-03	5.07E-03	5.70E-03	9
21	1	1	1	1E+00	0.5	32	100	3.22E-02	2.94E-02	2.61E-02	2.93E-02	77
22	4	20	14	1E+00	0	32	1213	2.36E-02	2.39E-02	2.55E-02	2.43E-02	66
23	2	15	8	3E-04	0	16	100	6.61E-03	6.22E-03	6.57E-03	6.45E-03	11
24	5	1	20	8E-05	0.2	128	1500	1.79E-02	8.86E-03	1.43E-02	1.37E-02	39
25	3	20	20	6E-02	0	64	526	1.63E-02	7.72E-03	1.58E-02	1.32E-02	38
26	1	20	10	9E-04	0.5	128	100	6.88E-03	9.33E-03	6.97E-03	7.75E-03	17
27	6	1	20	6E-02	0.2	32	100	2.34E-02	2.33E-02	2.32E-02	2.33E-02	61
28	2	20	1	1E+00	0.3	32	288	3.31E-02	3.02E-02	3.10E-02	3.15E-02	82

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
29	4	20	1	2E-04	0.3	32	1500	2.31E-02	2.24E-02	2.36E-02	2.30E-02	58
30	5	11	11	2E-03	0	128	100	4.78E-03	4.77E-03	4.43E-03	4.66E-03	2
31	5	14	20	1E+00	0.3	64	159	9.05E-01	4.18E+00	1.22E+00	2.10E+00	99
32	5	18	20	4E-02	0	64	633	9.72E-03	1.38E-02	9.38E-03	1.10E-02	32
33	6	20	18	2E-04	0.3	64	1231	9.17E-03	1.03E-02	1.05E-02	1.00E-02	25
34	2	18	16	1E-06	0	128	100	1.42E-01	1.25E-01	2.36E-01	1.68E-01	96
35	6	18	6	9E-02	0	32	100	1.99E-02	2.27E-02	2.45E-02	2.24E-02	55
36	6	18	20	8E-02	0	64	100	1.61E-02	1.58E-02	1.79E-02	1.65E-02	46
37	3	20	1	6E-03	0.3	32	893	2.35E-02	1.83E-02	2.38E-02	2.18E-02	53
38	5	12	20	4E-05	0.1	64	305	1.62E-02	1.21E-02	1.42E-02	1.42E-02	42
39	1	20	20	1E-06	0	128	100	2.97E-01	1.16E-01	1.60E-01	1.91E-01	97
40	6	12	4	2E-02	0.1	16	211	1.62E-02	1.34E-02	1.94E-02	1.63E-02	45
41	5	16	20	8E-06	0	16	100	2.38E-02	2.23E-02	2.88E-02	2.50E-02	68
42	5	15	16	4E-05	0	64	305	9.88E-03	8.12E-03	8.05E-03	8.69E-03	20
43	6	12	18	2E-06	0.1	16	213	3.66E-02	3.71E-02	3.55E-02	3.65E-02	86

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
44	5	11	5	1E-03	0	32	591	7.99E-03	5.23E-03	6.52E-03	6.60E-03	13
45	1	20	1	3E-06	0.5	128	100	1.15E-01	1.64E-01	1.04E-01	1.27E-01	94
46	4	15	8	1E+00	0	32	1468	2.47E-02	2.56E-02	2.91E-02	2.65E-02	72
47	3	15	6	6E-04	0.3	32	1074	7.29E-03	7.48E-03	7.41E-03	7.39E-03	15
48	1	20	1	1E-02	0.5	128	100	5.46E-03	4.82E-03	4.20E-03	4.85E-03	3
49	1	15	17	1E-06	0.3	32	861	1.56E-01	1.39E-01	8.82E-02	1.28E-01	95
50	5	19	20	1E-06	0.3	64	206	3.96E-02	6.21E-02	4.84E-02	5.00E-02	90
51	1	15	8	3E-03	0.3	32	100	5.57E-03	4.84E-03	6.12E-03	5.49E-03	8
52	6	2	7	3E-04	0.2	64	586	1.39E-02	1.46E-02	1.50E-02	1.46E-02	44
53	3	7	19	5E-03	0.4	32	565	6.92E-03	9.01E-03	8.28E-03	8.02E-03	18
54	4	6	18	9E-03	0.4	64	1424	1.11E-02	1.10E-02	9.51E-03	1.05E-02	30
55	5	7	19	6E-03	0.4	128	1099	1.31E-02	1.31E-02	1.52E-02	1.38E-02	40
56	1	7	15	9E-06	0.4	16	135	1.67E-01	5.86E-01	1.19E-01	2.92E-01	98
57	1	9	11	8E-01	0.4	16	209	2.99E-02	2.60E-02	3.48E-02	3.02E-02	80
58	1	7	20	3E-02	0.4	16	100	9.52E-03	6.86E-03	6.36E-03	7.52E-03	16

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
59	5	16	20	1E-03	0	32	175	5.01E-03	5.24E-03	4.65E-03	4.97E-03	5
60	2	8	10	5E-02	0.4	32	407	9.84E-03	1.03E-02	8.95E-03	9.68E-03	24
61	3	5	14	2E-02	0.4	16	1277	9.01E-03	9.01E-03	1.25E-02	1.02E-02	28
62	4	4	11	1E-02	0.4	64	823	1.07E-02	1.24E-02	1.41E-02	1.24E-02	36
63	3	5	17	1E-02	0.4	64	781	8.03E-03	8.29E-03	1.47E-02	1.04E-02	29
64	5	4	7	2E-02	0.4	32	283	2.33E-02	2.39E-02	2.34E-02	2.35E-02	63
65	1	4	20	9E-04	0.4	128	1500	9.03E-03	5.75E-03	6.62E-03	7.11E-03	14
66	3	5	2	9E-01	0.4	16	836	2.66E-02	3.39E-02	3.06E-02	3.04E-02	81
67	4	5	1	8E-06	0.4	16	1500	2.46E-02	2.43E-02	2.47E-02	2.43E-02	67
68	4	6	13	3E-02	0.4	16	921	2.23E-02	2.36E-02	2.29E-02	2.29E-02	57
69	3	6	19	1E-04	0.4	64	770	8.05E-03	8.12E-03	1.04E-02	8.90E-03	21
70	4	8	18	2E-02	0.4	16	690	2.06E-02	2.33E-02	2.20E-02	2.19E-02	54
71	4	10	20	2E-03	0	64	216	5.35E-03	4.80E-03	4.98E-03	5.06E-03	6
72	2	8	1	5E-04	0.4	128	1071	1.96E-02	1.70E-02	1.82E-02	1.83E-02	50
73	2	10	18	2E-03	0	64	1264	1.11E-02	6.63E-03	1.25E-02	1.01E-02	27

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
74	6	10	20	7E-06	0	16	100	2.96E-02	3.12E-02	2.81E-02	2.96E-02	78
75	6	12	20	1E+00	0.1	16	678	6.69E+01	4.07E-02	2.50E-02	2.25E+01	100
76	5	9	4	9E-03	0.1	32	193	1.77E-02	1.06E-02	9.75E-03	1.26E-02	37
77	5	11	4	1E-02	0.1	32	516	1.18E-02	1.15E-02	1.23E-02	1.19E-02	35
78	6	1	1	6E-06	0.4	32	1500	3.34E-02	3.32E-02	3.26E-02	3.29E-02	85
79	1	15	2	1E+00	0.3	32	1068	2.76E-02	2.63E-02	2.83E-02	2.74E-02	74
80	1	12	1	3E-03	0.3	32	845	1.18E-02	5.58E-03	9.42E-03	8.93E-03	22
81	4	1	1	5E-05	0.4	32	1470	2.36E-02	2.32E-02	2.25E-02	2.31E-02	60
82	1	16	18	1E-03	0.3	32	235	4.42E-03	4.69E-03	4.69E-03	4.60E-03	1
83	6	16	3	1E-02	0.1	16	100	1.79E-02	1.69E-02	1.48E-02	1.65E-02	47
84	5	1	5	1E-02	0.1	32	241	1.41E-02	1.27E-02	1.58E-02	1.42E-02	41
85	6	1	19	4E-03	0.1	64	1500	1.38E-02	1.68E-02	1.24E-02	1.44E-02	43
86	6	10	20	2E-04	0.4	128	100	2.53E-02	2.49E-02	2.67E-02	2.57E-02	70
87	6	1	20	1E-04	0.1	16	100	1.21E-02	2.05E-02	1.93E-02	1.73E-02	49
88	1	4	14	3E-03	0.3	16	1500	1.02E-02	6.11E-03	9.26E-03	8.50E-03	19

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
89	3	1	1	1E-03	0.1	128	1474	2.28E-02	2.22E-02	2.03E-02	2.17E-02	52
90	6	19	20	8E-05	0.3	128	100	2.80E-02	2.62E-02	2.59E-02	2.66E-02	73
91	1	11	2	7E-03	0.4	128	1138	5.71E-03	9.48E-03	1.68E-02	1.07E-02	31
92	3	20	1	2E-05	0	32	100	4.54E-02	4.86E-02	4.51E-02	4.64E-02	89
93	1	1	20	9E-02	0.4	16	100	1.97E-02	2.58E-02	2.27E-02	2.27E-02	56
94	1	5	1	1E+00	0.3	32	1500	3.79E-02	2.96E-02	3.00E-02	3.24E-02	83
95	2	11	16	1E+00	0.3	16	901	3.10E-02	3.75E-02	4.18E-02	3.68E-02	88
96	2	7	5	1E+00	0.3	16	216	3.23E-02	2.98E-02	2.89E-02	3.01E-02	79
97	2	14	8	1E+00	0.3	16	952	3.15E-02	2.82E-02	4.99E-02	3.66E-02	87
98	6	4	1	7E-06	0.4	16	1500	2.58E-02	2.56E-02	2.53E-02	2.56E-02	69
99	4	9	2	3E-05	0.4	32	592	2.47E-02	2.26E-02	2.49E-02	2.41E-02	65
100	1	7	4	1E+00	0.3	32	458	2.85E-02	2.81E-02	2.83E-02	2.83E-02	76

Anexo 10. Modelos entrenados para predecir el CBR al 95 % de la máxima densidad seca utilizando 10 variables de entrada

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
1	5	13	10	8E-02	0.2	128	283	1.44E-02	1.54E-02	1.51E-02	1.50E-02	45
2	5	7	18	6E-04	0	128	832	5.26E-03	7.41E-03	4.80E-03	5.85E-03	23
3	1	10	12	1E-05	0.4	64	843	5.55E-02	6.81E-02	4.28E-02	5.55E-02	90
4	5	15	7	3E-06	0.3	64	560	2.54E-02	3.22E-02	2.47E-02	2.74E-02	80
5	1	11	19	2E-05	0.3	64	1193	1.96E-02	1.02E-02	2.24E-02	1.73E-02	60
6	3	6	5	9E-06	0.1	32	550	1.88E-02	3.59E-02	4.26E-02	3.23E-02	85
7	2	3	4	4E-01	0.2	64	195	1.66E-02	1.68E-02	1.58E-02	1.64E-02	56
8	2	12	11	2E-06	0.2	128	308	1.21E-01	8.84E-02	9.13E-02	1.00E-01	94
9	5	16	9	3E-04	0.1	32	1106	5.10E-03	4.23E-03	4.71E-03	4.69E-03	16
10	2	15	5	5E-03	0.3	16	441	6.27E-03	6.39E-03	4.66E-03	5.75E-03	21
11	2	5	13	7E-02	0	64	1500	1.38E-02	7.92E-03	8.69E-03	1.02E-02	31
12	5	2	20	4E-01	0	16	653	1.86E-02	1.69E-02	1.76E-02	1.77E-02	63
13	1	19	20	1E-04	0.1	16	1500	3.68E-03	4.78E-03	3.79E-03	4.03E-03	2

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
14	4	6	20	3E-04	0.1	64	1500	5.03E-03	4.35E-03	5.87E-03	5.06E-03	19
15	2	9	20	4E-05	0.1	32	1500	4.07E-03	4.18E-03	3.91E-03	4.06E-03	3
16	1	14	20	1E-06	0.4	16	819	2.09E-01	8.69E-02	9.52E-02	1.31E-01	97
17	6	3	6	3E-05	0.5	128	446	2.10E-02	2.25E-02	2.13E-02	2.16E-02	70
18	6	17	6	6E-03	0.1	32	1478	5.94E-03	6.77E-03	6.35E-03	6.36E-03	26
19	1	2	11	9E-04	0.4	128	1491	9.47E-03	7.17E-03	9.14E-03	8.60E-03	29
20	4	1	1	3E-03	0.5	16	155	1.54E-02	1.53E-02	1.54E-02	1.55E-02	47
21	1	20	1	2E-02	0.2	128	724	4.55E-03	4.42E-03	4.27E-03	4.40E-03	10
22	1	20	20	1E+00	0	16	1500	1.89E-02	1.71E-02	1.95E-02	1.84E-02	65
23	3	2	1	9E-05	0.5	128	100	4.31E-02	7.53E-02	3.92E-02	5.27E-02	89
24	6	17	2	1E-05	0	32	1021	1.61E-02	1.55E-02	1.68E-02	1.62E-02	55
25	3	12	13	8E-03	0.5	16	1402	5.78E-03	7.04E-03	6.17E-03	6.36E-03	25
26	6	20	20	7E-04	0	16	1500	6.49E-03	5.94E-03	5.04E-03	5.81E-03	22
27	3	19	1	6E-03	0	16	870	1.43E-02	1.46E-02	1.39E-02	1.43E-02	42
28	6	8	1	1E+00	0.5	32	1464	1.77E-02	2.00E-02	1.68E-02	1.81E-02	64

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
29	6	8	20	6E-05	0	32	431	3.82E-03	4.23E-03	4.87E-03	4.29E-03	9
30	6	19	10	1E-03	0.5	16	1500	1.19E-02	1.27E-02	1.17E-02	1.22E-02	35
31	1	20	20	5E-03	0.5	16	100	3.41E-03	4.46E-03	4.52E-03	4.14E-03	5
32	6	15	20	2E-02	0.1	16	666	1.22E-02	1.16E-02	1.11E-02	1.17E-02	33
33	1	1	20	7E-03	0.5	64	1500	1.15E-02	1.45E-02	1.35E-02	1.32E-02	37
34	1	16	12	2E-02	0	128	100	5.37E-03	5.76E-03	3.91E-03	5.00E-03	18
35	6	14	20	1E-04	0	128	1500	3.79E-03	4.94E-03	4.18E-03	4.29E-03	8
36	6	10	19	6E-02	0.5	64	1500	1.54E-02	1.53E-02	1.57E-02	1.55E-02	50
37	2	7	2	3E-02	0.5	32	100	1.45E-02	1.34E-02	1.53E-02	1.44E-02	43
38	6	11	7	5E-04	0	16	100	4.67E-03	5.64E-03	3.64E-03	4.65E-03	14
39	1	15	20	4E-05	0	16	1500	4.87E-03	4.75E-03	4.16E-03	4.62E-03	13
40	5	20	20	2E-05	0	32	632	4.66E-03	5.17E-03	5.62E-03	5.16E-03	20
41	1	20	20	1E-03	0	32	510	4.46E-03	3.95E-03	4.15E-03	4.17E-03	6
42	6	1	20	7E-06	0	32	545	1.56E-02	1.29E-02	1.35E-02	1.41E-02	40
43	1	20	20	2E-01	0.2	128	251	1.23E-02	1.20E-02	1.04E-02	1.16E-02	32

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
44	1	20	10	3E-03	0.4	16	100	4.03E-03	4.06E-03	4.21E-03	4.12E-03	4
45	4	8	20	5E-05	0	128	1076	4.89E-03	4.45E-03	4.35E-03	4.54E-03	12
46	6	20	1	1E+00	0.5	128	100	1.57E-02	1.54E-02	1.54E-02	1.55E-02	49
47	3	20	20	3E-03	0.1	64	100	3.77E-03	4.05E-03	4.01E-03	3.95E-03	1
48	1	20	1	1E-01	0.1	128	1407	9.97E-03	1.31E-02	1.21E-02	1.17E-02	34
49	6	1	20	3E-04	0.4	16	100	1.43E-02	1.36E-02	1.29E-02	1.36E-02	38
50	2	20	20	6E-04	0.5	64	293	4.65E-03	4.66E-03	4.77E-03	4.70E-03	17
51	3	20	11	9E-04	0	16	1500	4.98E-03	7.05E-03	5.74E-03	5.93E-03	24
52	3	20	20	3E-02	0.5	128	258	6.73E-03	7.95E-03	8.90E-03	7.85E-03	27
53	1	20	20	4E-04	0	128	100	2.02E-02	1.02E-02	1.30E-02	1.45E-02	44
54	6	20	1	2E-02	0.5	16	1449	1.56E-02	1.56E-02	1.55E-02	1.55E-02	51
55	6	20	20	4E-03	0	16	1020	2.91E-03	4.58E-03	5.31E-03	4.25E-03	7
56	1	20	1	1E+00	0.5	128	100	1.43E-02	1.28E-02	1.53E-02	1.41E-02	41
57	2	20	20	2E-03	0.5	16	1500	4.89E-03	4.02E-03	4.67E-03	4.52E-03	11
58	6	17	1	1E-06	0	128	1457	2.86E-02	6.38E-02	2.88E-02	4.04E-02	88

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
59	1	20	1	2E-03	0	32	1500	1.17E-02	7.74E-03	2.22E-02	1.39E-02	39
60	6	20	20	1E+00	0.5	128	100	8.17E+00	8.18E+00	5.91E+00	7.41E+00	100
61	6	1	1	1E-06	0	16	1500	2.53E-02	2.51E-02	2.52E-02	2.52E-02	78
62	1	1	20	2E-01	0.5	16	100	1.55E-02	1.75E-02	1.99E-02	1.76E-02	62
63	6	1	1	1E+00	0	16	727	2.89E-02	2.18E-02	2.19E-02	2.41E-02	76
64	6	1	20	1E-06	0.5	128	1500	2.59E-02	2.88E-02	2.65E-02	2.71E-02	79
65	5	20	12	1E+00	0	128	1500	2.49E-01	5.56E-02	1.77E-02	1.08E-01	95
66	5	20	1	1E+00	0	128	663	1.85E-02	2.01E-02	1.82E-02	1.89E-02	66
67	4	9	1	7E-06	0	64	1234	2.48E-02	2.21E-02	2.29E-02	2.32E-02	74
68	6	1	20	1E-06	0	16	100	2.73E-02	2.76E-02	2.97E-02	2.82E-02	82
69	4	1	20	1E+00	0.5	16	100	1.56E-02	1.62E-02	1.84E-02	1.68E-02	58
70	3	17	1	1E+00	0.2	64	1468	2.08E-02	2.21E-02	1.83E-02	2.03E-02	68
71	4	20	1	1E+00	0.5	16	100	2.15E-02	1.96E-02	2.16E-02	2.10E-02	69
72	5	3	16	9E-06	0.5	64	100	2.85E-02	3.35E-02	2.79E-02	2.99E-02	84
73	4	18	20	1E-06	0.3	64	1449	3.58E-02	2.54E-02	2.52E-02	2.87E-02	83

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
74	4	1	11	1E+00	0	16	1500	1.90E-02	1.89E-02	1.98E-02	1.93E-02	67
75	4	12	12	6E-03	0.4	128	1500	7.86E-03	7.79E-03	8.26E-03	8.00E-03	28
76	5	17	15	3E-01	0.4	32	638	1.64E-02	1.67E-02	1.64E-02	1.65E-02	57
77	3	11	16	1E+00	0.3	64	1500	1.64E-02	1.68E-02	1.95E-02	1.75E-02	61
78	4	15	2	1E-03	0.1	32	1500	1.27E-02	1.06E-02	1.49E-02	1.28E-02	36
79	3	20	18	1E+00	0	128	1500	2.58E-02	3.37E-02	1.53E-02	2.51E-02	77
80	5	20	17	1E-02	0	128	1500	4.18E-03	4.19E-03	5.64E-03	4.66E-03	15
81	4	15	12	1E-06	0.5	64	100	4.19E-02	2.49E-02	4.33E-02	3.68E-02	87
82	5	17	1	1E+00	0.5	16	100	2.04E-02	2.85E-02	1.80E-02	2.23E-02	73
83	2	20	20	1E-06	0.3	16	100	8.02E-02	8.02E-02	1.07E-01	8.93E-02	92
84	6	1	1	1E+00	0.5	16	100	2.03E-02	2.11E-02	2.87E-02	2.34E-02	75
85	4	16	13	1E+00	0	128	781	1.56E-02	1.62E-02	1.67E-02	1.61E-02	54
86	2	18	1	1E-06	0.2	128	1500	7.28E-02	4.66E-02	6.78E-02	6.28E-02	91
87	1	1	1	1E-06	0.5	16	1500	1.48E-01	1.45E-01	1.38E-01	1.42E-01	99
88	4	11	1	4E-02	0.3	128	100	1.55E-02	1.51E-02	1.55E-02	1.53E-02	46

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
89	1	20	16	1E-06	0.4	16	100	1.50E-01	1.04E-01	1.40E-01	1.32E-01	98
90	3	20	17	7E-06	0	128	388	2.35E-02	3.84E-02	2.05E-02	2.76E-02	81
91	2	19	1	1E-06	0.4	32	1500	3.26E-02	2.31E-01	6.67E-02	1.11E-01	96
92	1	1	1	1E+00	0.5	16	803	2.27E-02	2.51E-02	1.86E-02	2.20E-02	71
93	6	12	1	1E+00	0.5	16	226	2.13E-02	2.00E-02	2.51E-02	2.20E-02	72
94	6	7	5	7E-05	0.4	128	487	1.68E-02	1.77E-02	1.73E-02	1.73E-02	59
95	6	1	10	1E+00	0.5	32	100	1.54E-02	1.56E-02	1.54E-02	1.55E-02	48
96	1	18	11	1E-06	0.5	32	1500	1.57E-01	5.05E-02	7.25E-02	9.34E-02	93
97	5	1	14	1E+00	0	64	860	2.45E-02	1.54E-02	6.35E-02	3.41E-02	86
98	5	15	16	6E-01	0.5	32	100	1.57E-02	1.56E-02	1.55E-02	1.56E-02	53
99	6	11	12	4E-02	0.5	128	1500	1.55E-02	1.55E-02	1.58E-02	1.56E-02	52
100	2	6	1	9E-05	0.1	32	1141	6.49E-03	1.03E-02	1.08E-02	9.16E-03	30

Anexo 11. Modelos entrenados para predecir la máxima densidad seca utilizando 8 variables de entrada

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
1	1	8	3	1E-05	0	32	1292	6.10E-02	9.44E-02	8.89E-02	8.15E-02	85
2	2	6	12	2E-02	0.1	32	532	2.36E-02	2.09E-02	3.14E-02	2.54E-02	41
3	3	16	14	1E-01	0.1	32	160	2.50E-02	3.01E-02	1.04E-01	5.32E-02	83
4	5	13	18	1E-05	0.4	64	122	2.44E-01	2.33E-01	2.01E-01	2.26E-01	88
5	1	16	7	2E-04	0.4	16	572	1.11E-02	1.11E-02	1.13E-02	1.12E-02	6
6	2	5	15	2E-03	0.5	16	295	1.28E-02	1.35E-02	1.21E-02	1.28E-02	13
7	1	13	11	3E-04	0.2	32	664	1.07E-02	1.12E-02	1.22E-02	1.14E-02	8
8	2	5	3	4E-02	0.1	16	236	2.17E-02	1.64E-02	1.78E-02	1.86E-02	29
9	3	13	12	3E-04	0	32	1442	1.02E-02	1.12E-02	1.10E-02	1.08E-02	5
10	4	15	18	3E-02	0	64	1218	9.94E-03	7.36E-03	8.36E-03	8.55E-03	1
11	6	18	20	1E-06	0.3	16	1500	1.11E-01	1.02E-01	1.13E-01	1.09E-01	86
12	4	1	1	1E-06	0.5	128	1500	2.86E-01	3.32E-01	2.87E-01	3.02E-01	95
13	1	1	1	2E-01	0	128	1500	2.95E-02	2.59E-02	3.04E-02	2.86E-02	49
14	6	20	13	1E+00	0.5	64	1500	3.18E-02	3.18E-02	3.17E-02	3.18E-02	55

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
15	2	11	12	1E+00	0.4	16	449	3.64E-02	3.51E-02	4.27E-02	3.81E-02	68
16	1	20	1	6E-04	0.5	32	100	1.40E-02	1.44E-02	1.69E-02	1.51E-02	21
17	1	1	20	6E-04	0	32	1500	2.23E-02	1.86E-02	2.54E-02	2.21E-02	38
18	6	20	1	9E-03	0	128	1500	3.19E-02	3.20E-02	3.18E-02	3.19E-02	58
19	6	20	1	7E-03	0	16	1500	3.19E-02	3.18E-02	3.20E-02	3.19E-02	59
20	6	1	2	5E-02	0	64	1500	2.98E-02	3.13E-02	2.91E-02	3.00E-02	51
21	5	20	19	8E-02	0	128	1500	2.11E-02	1.61E-02	2.59E-02	2.11E-02	34
22	6	20	14	1E-02	0	64	1500	8.38E-03	1.33E-02	1.22E-02	1.13E-02	7
23	5	12	9	3E-04	0	32	1500	9.69E-03	9.52E-03	1.24E-02	1.06E-02	4
24	6	1	1	1E-04	0	16	1500	3.21E-02	3.19E-02	3.19E-02	3.20E-02	60
25	6	20	3	3E-04	0	32	1422	1.51E-02	2.05E-02	1.99E-02	1.84E-02	27
26	6	20	20	1E+00	0.5	128	100	7.43E+01	5.85E+01	1.95E+01	5.10E+01	100
27	6	20	1	1E-06	0	64	598	2.94E-01	2.96E-01	2.81E-01	2.90E-01	94
28	6	1	10	1E-06	0.5	16	145	2.97E-01	2.78E-01	2.77E-01	2.84E-01	93
29	1	7	5	4E-01	0.1	16	311	3.73E-02	3.59E-02	4.30E-02	3.87E-02	70

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
30	1	1	1	1E+00	0.2	16	1473	4.11E-02	5.32E-02	5.37E-02	4.93E-02	82
31	1	20	20	1E-06	0.5	32	100	4.21E-01	2.28E-01	3.74E-01	3.41E-01	96
32	2	8	1	1E+00	0.1	32	541	3.70E-02	3.49E-02	3.87E-02	3.69E-02	67
33	3	16	1	1E+00	0.1	32	157	4.03E-02	3.74E-02	3.90E-02	3.89E-02	71
34	2	17	1	1E+00	0.2	32	675	3.64E-02	5.40E-02	3.75E-02	4.28E-02	77
35	2	7	1	1E+00	0	32	235	4.05E-02	4.87E-02	5.66E-02	4.85E-02	80
36	1	9	1	4E-01	0	16	366	3.57E-02	3.87E-02	3.46E-02	3.64E-02	65
37	6	20	20	1E-06	0	32	1500	1.50E-01	1.22E-01	1.63E-01	1.45E-01	87
38	1	1	1	1E+00	0	32	1453	3.91E-02	3.69E-02	4.14E-02	3.91E-02	72
39	6	5	1	1E+00	0	32	613	4.40E-02	3.86E-02	3.22E-02	3.82E-02	69
40	3	11	1	2E-03	0.4	32	658	2.79E-02	2.88E-02	2.94E-02	2.88E-02	50
41	4	20	20	1E-03	0	64	1335	1.08E-02	9.68E-03	1.95E-02	1.33E-02	15
42	1	20	1	1E+00	0.4	32	578	3.48E-02	3.92E-02	3.59E-02	3.67E-02	66
43	1	16	6	6E-03	0	32	678	9.66E-03	1.10E-02	3.46E-02	1.85E-02	28
44	1	5	1	1E+00	0.5	32	100	2.98E-02	3.25E-02	3.51E-02	3.24E-02	62

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
45	1	20	1	2E-02	0.1	32	543	2.80E-02	4.12E-02	1.33E-02	2.77E-02	46
46	1	13	1	1E+00	0.5	16	597	4.25E-02	4.11E-02	4.22E-02	4.19E-02	76
47	1	20	20	1E-06	0	16	349	3.92E-01	3.99E-01	4.10E-01	4.00E-01	97
48	1	1	1	1E+00	0.5	16	334	4.13E-02	5.91E-02	4.60E-02	4.87E-02	81
49	2	13	5	7E-03	0.5	16	598	1.46E-02	1.39E-02	1.88E-02	1.57E-02	23
50	2	3	1	1E+00	0.5	16	582	4.79E-02	3.97E-02	5.19E-02	4.64E-02	79
51	3	1	10	2E-02	0.5	16	182	2.92E-02	3.13E-02	3.00E-02	3.02E-02	52
52	1	8	1	8E-03	0.5	16	580	2.77E-02	1.31E-02	3.46E-02	2.52E-02	39
53	6	1	9	3E-02	0.5	16	308	3.22E-02	3.17E-02	3.15E-02	3.19E-02	57
54	5	13	7	8E-03	0	32	1470	1.63E-02	1.11E-02	1.09E-02	1.27E-02	12
55	1	20	20	8E-03	0.5	16	641	1.55E-02	8.91E-03	1.05E-02	1.17E-02	11
56	4	18	20	1E+00	0	64	1259	4.01E-02	4.83E-02	9.74E-02	6.21E-02	84
57	5	18	18	5E-04	0.3	64	1198	1.24E-02	1.38E-02	1.40E-02	1.34E-02	16
58	1	20	20	2E-03	0.5	64	1500	1.11E-02	2.15E-02	1.57E-02	1.60E-02	25
59	2	17	1	2E-02	0.4	32	101	2.66E-02	2.30E-02	2.84E-02	2.60E-02	42

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
60	4	1	1	6E-03	0	32	1481	3.17E-02	3.18E-02	3.20E-02	3.19E-02	56
61	1	8	1	1E+00	0.5	16	629	4.31E-02	4.84E-02	4.23E-02	4.45E-02	78
62	6	20	15	1E-06	0.1	64	1500	2.27E-01	2.38E-01	2.33E-01	2.33E-01	89
63	4	18	16	6E-03	0	32	1482	1.15E-02	9.77E-03	9.58E-03	1.03E-02	3
64	1	20	20	4E-02	0.1	64	1500	3.61E-02	1.20E-02	1.68E-02	2.14E-02	35
65	2	12	17	1E+00	0.3	64	1226	3.37E-02	3.25E-02	3.26E-02	3.30E-02	64
66	3	1	5	1E+00	0.5	16	236	3.71E-02	4.47E-02	3.73E-02	3.98E-02	73
67	3	16	16	1E-02	0.5	64	1236	1.68E-02	1.48E-02	1.53E-02	1.57E-02	22
68	1	11	12	1E+00	0.5	64	1399	4.07E-02	4.75E-02	3.68E-02	4.16E-02	75
69	1	1	20	8E-03	0.5	64	1157	2.07E-02	2.08E-02	3.42E-02	2.52E-02	40
70	2	19	3	1E-02	0.5	64	1369	1.71E-02	1.82E-02	1.94E-02	1.82E-02	26
71	1	19	13	8E-03	0	32	1429	1.20E-02	1.47E-02	1.38E-02	1.35E-02	18
72	1	20	1	2E-03	0.5	16	848	1.41E-02	3.77E-02	3.00E-02	2.74E-02	44
73	6	18	20	9E-03	0	32	1411	1.23E-02	7.72E-03	8.62E-03	9.56E-03	2
74	1	20	1	2E-02	0.5	16	478	1.31E-02	9.64E-03	3.55E-02	1.95E-02	31

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
75	1	1	20	8E-03	0.5	16	591	2.17E-02	2.08E-02	2.20E-02	2.15E-02	36
76	1	2	20	7E-03	0	32	786	1.75E-02	1.91E-02	2.01E-02	1.89E-02	30
77	5	18	20	8E-03	0.3	64	1220	1.60E-02	1.35E-02	1.84E-02	1.59E-02	24
78	1	20	18	1E+00	0	64	100	3.22E-02	2.91E-02	3.20E-02	3.10E-02	53
79	2	20	1	3E-02	0.5	32	133	2.75E-02	2.72E-02	2.67E-02	2.71E-02	43
80	3	20	20	1E-02	0.4	32	1498	1.38E-02	1.37E-02	1.30E-02	1.35E-02	17
81	6	9	19	2E-02	0.4	64	1423	2.74E-02	3.03E-02	2.58E-02	2.79E-02	47
82	1	20	20	5E-03	0	64	1155	9.44E-03	1.01E-02	1.44E-02	1.14E-02	9
83	6	20	1	1E+00	0	64	100	4.48E-02	4.06E-02	3.93E-02	4.15E-02	74
84	1	1	1	2E-03	0	16	1500	2.05E-02	2.57E-02	1.99E-02	2.20E-02	37
85	2	20	17	8E-03	0.2	64	1212	8.63E-03	2.38E-02	9.19E-03	1.39E-02	20
86	1	6	20	1E+00	0.1	64	100	3.28E-02	3.29E-02	3.30E-02	3.28E-02	63
87	2	20	20	9E-03	0.5	64	1225	1.25E-02	1.00E-02	1.20E-02	1.15E-02	10
88	4	20	11	1E-02	0.4	64	1166	2.01E-02	2.03E-02	1.90E-02	1.98E-02	33
89	6	1	1	1E-03	0.5	16	1500	3.21E-02	3.20E-02	3.20E-02	3.20E-02	61

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
90	1	1	1	1E-06	0	128	1500	5.53E-01	4.54E-01	3.31E-01	4.43E-01	98
91	1	1	1	4E-03	0	16	1315	2.54E-02	2.97E-02	2.70E-02	2.75E-02	45
92	3	7	1	6E-02	0	128	1500	2.73E-02	2.60E-02	3.07E-02	2.80E-02	48
93	6	11	5	2E-01	0.5	128	1500	3.15E-02	3.14E-02	3.14E-02	3.14E-02	54
94	6	1	20	1E-02	0	16	1500	1.96E-02	1.69E-02	2.24E-02	1.96E-02	32
95	6	20	20	1E+00	0.5	64	460	2.42E+00	2.37E+00	3.83E+00	2.89E+00	99
96	2	12	2	1E-06	0	64	966	2.83E-01	2.75E-01	2.67E-01	2.76E-01	91
97	3	12	15	1E-06	0.2	64	100	2.57E-01	2.72E-01	2.80E-01	2.69E-01	90
98	6	19	1	1E-06	0	64	672	2.81E-01	2.81E-01	2.82E-01	2.81E-01	92
99	1	14	14	1E-02	0	64	100	1.73E-02	1.28E-02	1.13E-02	1.38E-02	19
100	1	20	20	2E-02	0.5	16	270	1.07E-02	1.18E-02	1.63E-02	1.29E-02	14

Anexo 12. Modelos entrenados para predecir el óptimo contenido de humedad utilizando 8 variables de entrada

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
1	3	19	15	2E-03	0.3	64	493	1.44E-02	1.74E-02	1.49E-02	1.56E-02	6
2	4	20	20	3E-05	0	64	858	1.55E-02	1.69E-02	1.67E-02	1.64E-02	8
3	5	12	2	2E-05	0	64	403	1.88E-01	1.92E-01	2.05E-01	1.95E-01	84
4	2	10	17	9E-04	0.4	128	101	2.02E-02	1.99E-02	2.03E-02	2.01E-02	33
5	5	10	5	5E-06	0.5	32	656	1.74E-01	1.81E-01	1.86E-01	1.80E-01	81
6	2	7	18	3E-02	0.4	64	931	2.08E-02	1.75E-02	1.79E-02	1.87E-02	25
7	5	12	5	1E-01	0.4	16	558	3.51E-02	3.47E-02	3.35E-02	3.44E-02	58
8	6	18	11	9E-02	0	64	1126	2.64E-02	2.60E-02	2.53E-02	2.59E-02	45
9	4	17	8	2E-06	0.4	64	701	1.99E-01	1.81E-01	1.89E-01	1.90E-01	82
10	4	13	2	3E-01	0.3	64	240	3.42E-02	3.45E-02	3.81E-02	3.55E-02	60
11	6	1	20	5E-04	0	16	1500	2.02E-02	2.45E-02	2.07E-02	2.18E-02	37
12	4	20	20	1E-03	0.2	16	1053	1.64E-02	2.08E-02	1.48E-02	1.73E-02	18
13	6	20	20	1E+00	0.5	16	100	3.23E+00	5.25E+00	3.54E+00	3.99E+00	99
14	6	17	10	7E-02	0	64	1211	2.19E-02	3.07E-02	2.40E-02	2.55E-02	44

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
15	5	13	5	1E+00	0.3	16	489	4.13E-02	3.73E-02	4.36E-02	4.07E-02	65
16	5	7	1	1E-06	0.5	32	801	2.21E-01	2.21E-01	2.21E-01	2.21E-01	87
17	4	20	20	4E-04	0	32	959	2.04E-02	1.85E-02	1.87E-02	1.92E-02	28
18	1	17	1	1E-06	0.3	128	100	4.78E-01	1.83E-01	4.40E-01	3.67E-01	96
19	3	19	18	2E-02	0.3	64	537	1.75E-02	1.70E-02	1.70E-02	1.72E-02	15
20	3	19	12	2E-04	0.3	128	445	2.08E-02	2.13E-02	1.97E-02	2.06E-02	35
21	3	19	20	6E-01	0.3	16	542	3.45E-02	4.63E-02	4.14E-02	4.07E-02	66
22	5	17	20	3E-02	0.4	32	151	3.00E-02	2.68E-02	2.77E-02	2.82E-02	48
23	3	2	1	1E-06	0.4	128	205	2.25E-01	2.42E-01	2.28E-01	2.32E-01	88
24	4	18	19	1E-06	0.1	128	371	2.74E-01	2.09E-01	2.19E-01	2.34E-01	90
25	3	19	1	1E-06	0	128	427	2.65E-01	2.30E-01	2.67E-01	2.54E-01	93
26	3	20	1	1E-06	0.3	128	464	2.59E-01	2.30E-01	2.38E-01	2.42E-01	92
27	4	16	8	1E-06	0	128	100	2.04E-01	1.97E-01	2.21E-01	2.07E-01	86
28	4	20	20	4E-04	0.2	16	1226	1.55E-02	1.77E-02	1.60E-02	1.64E-02	9
29	4	19	7	1E-06	0.2	16	1500	1.11E-01	1.41E-01	1.32E-01	1.28E-01	78

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
30	4	20	20	5E-05	0	64	1500	1.54E-02	1.49E-02	1.49E-02	1.51E-02	2
31	4	3	20	5E-02	0.3	16	1220	3.35E-02	3.36E-02	3.34E-02	3.35E-02	56
32	2	1	20	1E+00	0.3	16	1500	3.99E-02	4.79E-02	4.05E-02	4.28E-02	68
33	2	16	1	4E-05	0.3	32	452	1.01E-01	7.93E-02	9.26E-02	9.09E-02	75
34	5	18	20	3E-02	0.3	64	921	2.51E-02	2.48E-02	3.00E-02	2.66E-02	46
35	3	3	16	1E-06	0.3	128	100	3.88E-01	2.57E-01	2.70E-01	3.05E-01	94
36	3	19	16	5E-04	0.3	128	149	1.93E-02	2.25E-02	1.99E-02	2.05E-02	34
37	3	19	16	1E-01	0.3	64	532	3.08E-02	3.35E-02	2.71E-02	3.04E-02	51
38	3	19	18	2E-05	0.3	128	470	1.07E-01	1.03E-01	6.65E-02	9.23E-02	76
39	5	20	20	1E+00	0.5	32	100	2.86E+00	6.19E+00	2.18E+00	3.74E+00	98
40	1	20	1	1E+00	0.1	32	1216	4.43E-02	4.19E-02	4.61E-02	4.41E-02	70
41	1	20	20	5E-06	0.5	128	100	4.52E-01	3.13E-01	2.79E-01	3.46E-01	95
42	4	20	1	1E-04	0.3	128	1159	7.99E-02	7.14E-02	7.72E-02	7.61E-02	74
43	6	13	20	2E-02	0	32	1007	1.84E-02	2.07E-02	1.74E-02	1.88E-02	26
44	2	20	20	4E-05	0	64	100	1.10E-01	8.55E-02	1.06E-01	1.00E-01	77

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
45	3	20	1	6E-03	0.4	16	1500	3.25E-02	3.23E-02	3.27E-02	3.25E-02	53
46	5	20	20	3E-04	0	32	1500	2.39E-02	2.16E-02	2.12E-02	2.23E-02	39
47	1	13	20	6E-04	0.2	64	1500	1.61E-02	1.67E-02	1.87E-02	1.72E-02	16
48	1	4	20	5E-03	0.4	128	714	2.01E-02	1.63E-02	1.69E-02	1.78E-02	20
49	2	20	20	2E-03	0.3	128	1500	1.83E-02	1.62E-02	1.75E-02	1.73E-02	17
50	6	1	5	1E+00	0	16	1500	4.65E-02	4.39E-02	4.89E-02	4.65E-02	73
51	3	13	20	7E-04	0.2	64	1430	1.50E-02	1.61E-02	1.87E-02	1.66E-02	10
52	2	3	20	7E-03	0.4	32	108	1.76E-02	1.87E-02	1.91E-02	1.85E-02	23
53	1	20	20	1E+00	0.4	128	1500	3.30E-02	3.30E-02	3.34E-02	3.31E-02	55
54	6	1	1	1E+00	0.1	16	682	3.83E-02	4.25E-02	4.80E-02	4.30E-02	69
55	3	20	20	7E-04	0.2	16	328	1.56E-02	1.74E-02	1.52E-02	1.61E-02	7
56	2	7	20	7E-04	0.2	32	906	1.52E-02	1.63E-02	1.86E-02	1.67E-02	11
57	5	1	20	7E-03	0	32	174	2.14E-02	1.87E-02	1.95E-02	1.99E-02	31
58	2	19	20	7E-04	0.2	32	188	1.44E-02	1.59E-02	1.43E-02	1.49E-02	1
59	3	20	20	6E-04	0.2	128	100	1.82E-02	1.74E-02	1.71E-02	1.76E-02	19

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
60	2	16	20	6E-04	0.2	16	1500	2.99E-02	2.37E-02	2.27E-02	2.55E-02	43
61	1	1	20	2E-03	0.2	128	100	4.19E-02	5.31E-02	4.32E-02	4.61E-02	71
62	1	1	20	3E-04	0.2	64	100	1.06E-01	2.20E-01	2.87E-01	2.03E-01	85
63	2	20	20	1E-03	0.2	32	1500	2.47E-02	2.36E-02	2.50E-02	2.44E-02	42
64	6	20	19	4E-03	0.2	16	1500	1.79E-02	2.32E-02	1.90E-02	2.00E-02	32
65	3	20	20	2E-03	0.2	64	1500	1.68E-02	2.23E-02	1.68E-02	1.87E-02	24
66	2	9	19	2E-03	0.2	32	248	1.59E-02	1.61E-02	1.46E-02	1.55E-02	4
67	6	20	15	7E-04	0.2	16	1500	1.92E-02	1.99E-02	3.00E-02	2.31E-02	40
68	6	20	18	4E-03	0.2	16	1500	2.33E-02	2.22E-02	2.09E-02	2.21E-02	38
69	6	8	20	4E-03	0.2	128	1396	1.85E-02	1.95E-02	1.94E-02	1.92E-02	27
70	6	20	5	1E-06	0.2	16	1500	2.03E-01	1.76E-01	1.54E-01	1.77E-01	80
71	6	20	20	2E-01	0	128	1500	3.34E-02	3.38E-02	3.34E-02	3.36E-02	57
72	5	20	9	4E-03	0.2	16	1500	2.37E-02	2.04E-02	2.65E-02	2.36E-02	41
73	1	20	20	3E-04	0.2	128	100	3.26E-02	3.71E-02	2.78E-02	3.25E-02	54
74	3	11	10	2E-03	0.2	128	1500	1.66E-02	1.49E-02	2.31E-02	1.82E-02	22

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
75	6	1	20	3E-03	0	16	1023	1.93E-02	1.85E-02	2.16E-02	1.98E-02	30
76	6	1	20	5E-02	0	32	675	3.08E-02	3.12E-02	2.76E-02	2.99E-02	50
77	5	20	20	3E-02	0	64	1085	1.75E-02	1.60E-02	1.69E-02	1.68E-02	12
78	1	14	20	6E-03	0.4	128	1101	1.82E-02	1.78E-02	2.93E-02	2.17E-02	36
79	1	1	20	6E-02	0.4	64	1194	2.98E-02	2.22E-02	3.33E-02	2.84E-02	49
80	3	1	13	1E+00	0.1	32	856	3.59E-02	4.28E-02	3.71E-02	3.86E-02	63
81	3	1	4	1E+00	0.1	64	1202	4.54E-02	3.49E-02	3.74E-02	3.93E-02	64
82	5	18	20	2E-02	0	16	500	1.70E-02	1.83E-02	1.60E-02	1.71E-02	14
83	2	20	20	4E-04	0.3	64	361	1.54E-02	1.52E-02	1.59E-02	1.56E-02	5
84	3	1	17	1E+00	0.1	32	836	4.04E-02	4.27E-02	3.92E-02	4.08E-02	67
85	3	8	20	7E-03	0.3	64	304	1.67E-02	1.80E-02	1.92E-02	1.80E-02	21
86	3	1	1	1E+00	0.1	16	747	4.44E-02	4.36E-02	5.11E-02	4.64E-02	72
87	6	20	20	1E+00	0.1	64	778	7.00E+00	9.93E+00	1.25E+01	9.86E+00	100
88	2	17	15	1E-06	0.1	128	684	1.66E-01	1.28E-01	2.78E-01	1.91E-01	83
89	1	1	20	4E-02	0.2	128	100	2.91E-02	2.54E-02	2.84E-02	2.76E-02	47

Modelo	Num layers	Units 1	Units 2	Learning rate	Drop rate	Batch size	Epochs	K1	K2	K3	MSE	Ranking
90	3	18	20	1E+00	0.3	128	157	6.08E-01	4.93E-01	7.02E-01	5.99E-01	97
91	6	20	18	1E+00	0.2	16	1309	3.61E-02	3.36E-02	3.45E-02	3.48E-02	59
92	1	18	1	1E-05	0.1	64	346	1.96E-01	8.37E-02	1.97E-01	1.59E-01	79
93	6	19	1	1E+00	0.3	64	100	3.76E-02	3.59E-02	4.02E-02	3.79E-02	62
94	5	20	20	3E-02	0.2	16	748	3.21E-02	3.12E-02	3.19E-02	3.17E-02	52
95	6	16	20	1E-06	0.1	128	164	2.53E-01	2.29E-01	2.19E-01	2.33E-01	89
96	3	19	4	1E+00	0.3	64	183	3.53E-02	3.91E-02	3.37E-02	3.60E-02	61
97	4	18	16	1E-06	0.1	32	251	2.17E-01	2.72E-01	2.31E-01	2.40E-01	91
98	4	18	18	8E-04	0.1	64	782	1.70E-02	1.53E-02	1.40E-02	1.54E-02	3
99	4	18	16	2E-03	0.1	16	475	1.77E-02	1.86E-02	2.14E-02	1.92E-02	29
100	4	18	17	2E-02	0.1	128	960	1.67E-02	1.68E-02	1.72E-02	1.69E-02	13