

Cálculo de la fracción volumétrica de las sustancias aceite y agua  
en flujo bifásico al interior de tuberías horizontales aplicando  
Long Short Term Memory (LSTM)

Alejandro Carreño Verdugo

Trabajo de Grado para Optar al Título de Ingeniero Mecánico

Director

Octavio Andrés González Estrada  
PhD en Ingeniería Mecánica y de Materiales

Codirector

Carlos Mauricio Ruiz Diaz  
Magister en Ingeniería Mecánica

Marlon Mauricio Hernández Cely  
PhD en Ingeniería Mecánica

Universidad Industrial de Santander  
Facultad de Ingenierías Fisicomecánicas  
Escuela de Ingeniería Mecánica  
Pregrado en Ingeniería Mecánica  
Bucaramanga

2023

### **Dedicatoria**

Dedico este trabajo de grado a mis padres, quienes han sido mi mayor fuente de inspiración y apoyo a lo largo de mi formación como Ingeniero Mecánico. Su sacrificio, aliento y amor incondicional constante me han impulsado a lograr mis objetivos y superar cada uno de los obstáculos que se me han presentado.

También dedico este trabajo a mi pareja, por su amor, paciencia y comprensión. Su apoyo constante ha sido una luz en mi camino, dándome motivación para seguir adelante en los momentos más desafiantes.

De igual manera lo dedico a mi familia, especialmente a mis abuelos, quienes siempre me han respaldado con gestos de afecto y palabras alentadoras en cada una de las etapas de mi formación.

### **Agradecimientos**

Agradezco primeramente a Dios por estar presente en cada día de mi vida, por ponerme en mi camino a personas que aportaron positivamente en este proceso de formación y así mismo por darme la oportunidad de ser un profesional.

Agradezco a mis profesores y asesores académicos, cuya sabiduría y orientación han sido fundamentales para mi desarrollo como Ingeniero Mecánico.

Agradezco al profesor Octavio Andrés González Estrada, director de mi proyecto de grado, por sus enseñanzas, por brindarme su orientación y ser guía en el desarrollo de este proyecto, por su paciencia y por haberme mostrado el camino de la investigación desde diferentes puntos de vista.

Agradezco a Carlos Mauricio Ruiz Diaz, codirector de este proyecto, por haberme acompañado en este proceso y darme la oportunidad de afrontar un reto como lo fue aprender sobre redes neuronales y aplicaciones de inteligencia artificial.

Agradezco a mis compañeros de estudio por su apoyo, colaboración y amistad a lo largo de esta travesía académica.

Expreso mi gratitud a quienes no he mencionado específicamente, pero que han dejado huella en mi camino como Ingeniero Mecánico. Cada encuentro, cada lección aprendida, cada experiencia compartida que ha contribuido a mi crecimiento personal y profesional.

**Tabla de Contenido**

	<b>Pág.</b>
Introducción .....	11
1. Objetivos .....	17
1.1 Objetivo General .....	17
1.2 Objetivos Específicos .....	17
2. Materiales y Métodos .....	18
2.1. Tratamiento de los Datos .....	19
2.2. Diseño de la Red Neuronal .....	19
3. Resultados.....	24
4. Conclusiones .....	38
5. Referencias .....	40

**Lista de Tablas**

	<b>Pág.</b>
Tabla 1 <i>Resumen Base de Datos Recopilada</i> .....	18
Tabla 2 <i>Parámetros del diseño de experimentos (DoE)</i> .....	24
Tabla 3 <i>Rendimiento del modelo variando los parámetros principales, usando 150 neuronas para su funcionamiento para las 10 configuraciones con el mejor MSE.</i> .....	25
Tabla 4 <i>Resultados rendimiento del modelo teniendo en cuenta una variación completa de parámetros principales, usando 100 neuronas para su funcionamiento</i> .....	26
Tabla 5 <i>Rendimiento del modelo variando los parámetros principales, usando 50 neuronas para su funcionamiento para las 10 configuraciones con el mejor MSE.</i> .....	27
Tabla 6 <i>Rendimiento del modelo variando los parámetros principales, usando 25 neuronas para su funcionamiento para las 10 configuraciones con el mejor MSE.</i> .....	28
Tabla 7 <i>Resultados generales de rendimiento teniendo como parámetro de referencia el MSE</i> ..	35
Tabla 8 <i>Selección de parámetros del sistema para obtener el mejor rendimiento del estudio</i> .....	36

**Lista de Figuras**

	<b>Pág.</b>
Figura 1 <i>Representación esquemática de la estructura de red neuronal artificial utilizada</i> .....	22
Figura 2 <i>Comportamiento de MSE relacionado con el número de neuronas y épocas del sistema</i> .....	29
Figura 3 <i>Correlación entre el número de neuronas en la capa oculta y el MSE del sistema</i> .....	30
Figura 4 <i>Variación del error absoluto medio (MAE) teniendo en cuenta el número de épocas del sistema</i> .....	31
Figura 5 <i>Variación del error absoluto medio porcentual (MAPE) teniendo en cuenta el número de neuronas del sistema</i> .....	32
Figura 6 <i>Análisis de comportamiento del error cuadrático medio (MSE), contra el número de épocas</i> .....	33
Figura 7 <i>Graficas de rendimiento de la configuración seleccionada</i> .....	36
Figura 8 <i>Desempeño del modelo seleccionado</i> .....	38

**Lista de Apéndices**

“Los apéndices están adjuntos y puede visualizarlos en la base de datos de la biblioteca UIS”

**Glosario (Opcional)**

D	Diámetro de la tubería
$\varepsilon_o$	Fracción volumétrica aceite
$\varepsilon_w$	Fracción volumétrica agua
$J_o$	Velocidad superficial del aceite
$J_{o+w}$	Velocidad superficial del flujo bifásico
$J_w$	Velocidad superficial del agua
$\mu_o$	Viscosidad del aceite
$\rho_o$	Densidad del aceite
LSTM	Long Short Term Memory
MAE	Error absoluto medio
MAPE	Error porcentual absoluto medio
MSE	Error cuadrático medio
RNN	Red neuronal recurrente



## Resumen

**Título:** Cálculo de la fracción volumétrica de las sustancias aceite y agua en flujo bifásico al interior de tuberías horizontales aplicando long short term memory (LSTM)\*

**Autor:** Alejandro Carreño Verdugo\*\*

**Palabras Clave:** LSTM, flujo bifásico líquido – líquido, tubería horizontal, redes neuronales recurrentes, industria 4.0.

**Descripción:** Este artículo presenta la implementación de una red neuronal Long Short Term Memory con el fin de calcular la fracción volumétrica del aceite y agua en un fluido bifásico que se transporta por una tubería horizontal. Se implementa una base de datos de la literatura que contiene pruebas de 13 autores para el entrenamiento del modelo, las entradas del modelo están dadas por un vector que contiene las velocidades superficiales de aceite y agua, la mezcla del fluido, el diámetro interno de la tubería y la viscosidad; teniendo como parámetro de salida el vector que contiene la fracción volumétrica del aceite. El proceso de entrenamiento y validación se realizan con una base de datos que contiene 2156 datos, los cuales se segmentan en 80% y 20% respectivamente, con el fin de obtener resultados concluyentes y poder realizar la validación del sistema. Se presentó un diseño de experimentos el cual permitía realizar todas las pruebas posibles al modelo según parámetros especificados anteriormente, habiendo realizado para el presente estudio 216 pruebas de rendimiento del modelo. Los resultados que presentó el modelo predictivo con mejor rendimiento evidencia tener un MSE de  $3,5651E-05$ , un MAE de 0.0045 y un MAPE de 3.0250%, este resultado se obtiene ingresando una función de transferencia ReLu, unas épocas de 20, una learnin rate de 0.1, función de transferencia sigmoid, un batch size de 1, optimizador ADAM y 150 neuronas en la capa oculta.

---

\* Trabajo de Grado

\*\* Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería Mecánica. Director: Octavio Andrés González Estrada. PhD en Ingeniería Mecánica y de Materiales. Codirector: Carlos Mauricio Ruiz Diaz, Magister en Ingeniería Mecánica. Marlon Mauricio Hernández Cely, PhD en Ingeniería Mecánica

### Abstract

**Title:** Calculation of the volume fraction of the substances oil and water in two-phase flow in horizontal pipes in two-phase flow inside horizontal pipelines by applying long short-term memory (LSTM)\*

**Author:** Alejandro Carreño Verdugo \*\*

**Key Words:** LSTM, Biphasic liquid-liquid flow, horizontal tubery, Recurrent neural network, Industry 4.0.

**Description:** This paper presents the implementation of a Long Short Term Memory neural network in order to calculate the volume fraction of oil and water in a two-phase fluid transported through a horizontal pipe. A literature database containing evidence from 13 authors is implemented for the training of the model, the model inputs are given by a vector containing the surface velocities of oil and water, the fluid mixture, the internal diameter of the pipe and the viscosity; having as output parameter the vector containing the oil volume fraction. The training and validation process is carried out with a database containing 2156 data, which are segmented into 80% and 20% respectively, in order to obtain conclusive results and be able to validate the system. A design of experiments was presented which allowed all possible tests to be performed on the model according to previously specified parameters, having carried out 216 model performance tests for the present study. The results presented by the predictive model with the best performance showed an MSE of 3.5651E-05, an MAE of 0.0045 and a MAPE of 3.0250%, this result was obtained by entering a ReLu transfer function, epochs of 20, a learnin rate of 0.1, sigmoid transfer function, a batch size of 1, ADAM optimizer and 150 neurons in the hidden layer.

---

\* Degree Work

\*\* Faculty of Physicomechanical Engineering. School of Mechanical Engineering. Director: Octavio Andrés González Estrada. PhD in Mechanical and Materials Engineering. Co-director: Carlos Mauricio Ruiz Diaz, Master in Mechanical Engineering. Marlon Mauricio Hernández Cely, PhD in Mechanical Engineering.

## Introducción

Los flujos multifásicos están presentes en procesos de la industria petroquímica y son de gran interés para las fases de producción y transporte del crudo y gas. Además, en busca de mejorar sus procesos, considerando el desarrollo de la industria 4.0, la industria petrolera ha centrado su interés en la implementación de tecnologías para la caracterización de flujos multifásicos (Bikmukhametov & Jäschke, 2020). Para realizar la clasificación de un flujo multifásico en tuberías, se requiere una descripción detallada de las diferentes distribuciones espaciales de las fases (líquidas o gaseosas), llamadas regímenes o patrones de flujo, las cuales hacen referencia a la configuración geométrica de los fluidos dentro de la tubería (Bikmukhametov & Jäschke, 2020). Predecir el comportamiento de los patrones de flujo, considerando las fases en las que se encuentren los fluidos, ya sea sólida, líquida o gas permite realizar estimaciones precisas de los parámetros de proceso, optimizar el uso de la infraestructura, del personal y reducir los gastos de operación (Bikmukhametov & Jäschke, 2020),(Consuegra et al., 2017).

Se han realizado muchos estudios del flujo bifásicos líquido – líquido para comprender su comportamiento en diferentes condiciones, especialmente, la mezcla aceite y agua (Osundare et al., 2020). Este tipo de flujo depende de numerosas variables como la geometría de la tubería (longitud, inclinación, material, diámetro), en igual forma como las propiedades presentes en la mezcla (tensión superficial, viscosidad y densidad) y los caudales de flujo (K. García, J. García, G. Montoya, M. T. Valecillos, n.d.). La estimación de los gradientes de presión y fracciones volumétricas del fluido se hace mediante la determinación de los patrones de flujo en el interior de una tubería (Osundare et al., 2020). La fracción volumétrica es una medida adimensional que

permite conocer la cantidad de cada componente de la mezcla en un sistema multifásico, el cual consiste en dividir el volumen del componente del sistema sobre el volumen total de la mezcla (Varotsis, 1997).

Adicionalmente, al momento de realizar explotación petrolera, cuantificar correctamente la fracción volumétrica es muy relevante para poder identificar la concentración de líquido y de gas en un pozo, facilitando la explotación de yacimientos de gas natural (Sattari et al., 2020). En una distribución líquido – líquido, como lo es una mezcla de agua y aceite, se puede utilizar para determinar la cantidad de agua y aceite que se encuentra en el interior de un pozo, e identificar las zonas con altas fracciones volumétricas del componente valioso que sea potencialmente extraíble (Malayeri et al., 2003).

En busca de una mayor precisión y flexibilidad en el procesamiento y análisis de datos de los procesos industriales en tiempo real, se han implementado mejoras en la conectividad a través del uso de tecnologías como el Internet de las cosas (IoT), el almacenamiento en la nube, la robótica y la inteligencia artificial (IA) (Abdelhafidh et al., 2018),(Slaný et al., 2020). Mediante la utilización de la inteligencia artificial (IA) y el aprendizaje autónomo, es posible realizar un análisis exhaustivo de grandes cantidades de datos y obtener predicciones precisas sobre las condiciones futuras de los procesos (Vuo et al., 2021). Como resultado, el análisis en tiempo real de la información recopilada por los dispositivos de medición y captación de imágenes en el estudio de los flujos multifásicos permite mejorar el proceso de modelado del sistema y comprender mejor los fenómenos presentes en el transporte de fluidos. De esta manera, se logra una mayor precisión en el modelado del funcionamiento del sistema y se mejora la comprensión de los fenómenos relacionados con el transporte de fluidos en el estudio de los flujos multifásicos (Ibrahim et al., 2018),(Sai et al., 2020).

La implementación de redes neuronales artificiales (RNA) es valiosa para dar solución a los problemas presentados en el control de procesos bifásicos. En efecto, juegan un papel importante en la separación de mezclas de líquidos, una RNA puede ser entrenada para variar el componente específico de un proceso, permitiendo así controlar la fracción volumétrica de este, con lo que se puede llegar a bajar costos dado que se optimiza el rendimiento del proceso (Ganeshmoorthy et al., 2019). Cabe agregar que el campo de acción de las RNA para hallar la fracción volumétrica en un sistema bifásico ha tenido un crecimiento notable en los últimos años (Özgür & Tosun, 2017). Esto se debe al desempeño presentado por las RNA para capacitarse en aprendizaje de patrones complejos de comportamiento en bases de datos de tamaño considerable, dándoles la capacidad de resolver problemas los cuales incurran en la caracterización de sistemas bifásicos (Chaki et al., 2014). Por su parte, las RNA tienen aplicaciones en el área de análisis de datos de espectroscopía, donde se usan para hallar la fracción volumétrica del sistema. Ahora bien, entre las ramas de la espectroscopía donde se aplican estas RNA se encuentra la espectroscopía infrarroja, la cual presenta datos sensibles a la presencia de algunos componentes en un sistema bifásico (Chaki et al., 2014). Por lo tanto, en el momento que se entrena un modelo compuesto por RNA, es posible predecir la fracción volumétrica de un componente específico del sistema. Además, como otra aplicación de las RNA resalta el análisis de bases de datos compuestas por imágenes, estos datos son obtenidos a través de la ayuda de cámaras y sensorica, así mismo pueden ser usados para determinar la distribución de fases en una distribución bifásica (Yoon & Kim, 2021). Como resultado, al entrenar un modelo con los suficientes datos es posible realizar la predicción de la fracción volumétrica de cada componente, recordando que la sumatoria de las mismas debe ser la unidad.

El aprendizaje automático es una rama de la inteligencia artificial que se enfoca en el desarrollo de modelos computacionales y algoritmos que permiten a las máquinas aprender a partir de la data y experiencias previas sin necesidad de programación específica (Cao et al., 2016). El proceso de aprendizaje consta de tres etapas: entrenamiento, validación y prueba. Durante la fase de entrenamiento, se alimenta el modelo con un conjunto de datos que incluyen muestras de entradas y salidas esperadas. Durante este proceso, el algoritmo ajusta los parámetros internos para minimizar el error entre las predicciones y las salidas esperadas utilizando un algoritmo optimizador (Cao et al., 2016). Posteriormente, se realiza la validación del modelo mediante la evaluación de un conjunto de datos que no se usó en el entrenamiento para comprobar que el modelo genere valores precisos en la salida y no esté sobreajustado con los datos de entrenamiento (Luo et al., 2018). Los modelos de aprendizaje automático pueden ser supervisados, no supervisados o de aprendizaje por refuerzo, dependiendo de los requisitos de la aplicación (Osarogiagbon et al., 2021). Los modelos supervisados se utilizan para predecir una salida a partir de datos de entrada, mientras que los modelos no supervisados se utilizan para identificar patrones en los datos sin requerir salidas previas (Zhong et al., 2019). El aprendizaje por refuerzo se centra en aprender la toma de decisiones a través de interacciones con el entorno (Osarogiagbon et al., 2021). El Deep Learning es una técnica de aprendizaje automático basada en redes neuronales con múltiples capas que permite procesar grandes conjuntos de datos sin estructuración clara, como imágenes, voz, texto y video, para identificar patrones y relaciones complejas que no son fáciles de detectar por humanos (Salgado et al., 2021). En la industria del petróleo y gas, el Deep Learning se utiliza para la optimización del proceso de producción, el mantenimiento predictivo, el monitoreo de equipos, la predicción de reservas de petróleo y gas y

el procesamiento y análisis de imágenes satelitales relacionadas con la exploración y producción petrolera (Temirchev et al., 2020).

En el procesamiento de secuencias de datos, como texto y audio, se emplean las redes neuronales recurrentes (RNN), un tipo de modelo de aprendizaje automático que tiene la capacidad de recordar información previamente trabajada y utilizarla para hacer predicciones futuras (Biswas, 2019). La arquitectura de las RNN se comporta de tal forma que la salida de una neurona se utiliza como entrada de la siguiente, lo que permite que la información fluya a través del tiempo y se capturen patrones de comportamiento en secuencias temporales complejas (Biswas, 2019). No obstante, el rendimiento de las RNN enfrenta dificultades en el entrenamiento de redes profundas debido a la desaparición del gradiente. Por lo tanto, se desarrolló una variante de las RNN llamada redes LSTM (Long Short Term Memory) que resuelve esta problemática (Madasu & Rangarajan, 2019). Las redes LSTM difieren de las RNN en su estructura, ya que funcionan como compuertas que controlan el flujo de información al decidir qué información es importante recordar y qué información no es relevante (Madasu & Rangarajan, 2019). Las redes LSTM se componen de células de memoria conectadas entre sí por compuertas de entrada, salida y olvido (Li et al., 2020). La función de estas compuertas es controlar el almacenamiento de información en cada célula de manera que la red aprenda a recordar solo la información relevante para el proceso y descarte la información poco importante (Sagheer & Kotb, 2019). Las redes LSTM han sido muy exitosas en diversas aplicaciones, incluyendo el procesamiento del lenguaje natural, la regresión de patrones de comportamiento y la clasificación de imágenes, gracias a su capacidad para trabajar con secuencias de datos largas y complejas de manera eficiente (Sagheer & Kotb, 2019).

En este trabajo se va a enfatizar en el estudio del flujo bifásico de aceite y agua, el cual está caracterizado a partir de la identificación de comportamientos característicos, la viscosidad de las sustancias, el diferencial de presión en la línea y fracciones volumétricas. El objetivo es investigar la viabilidad de la aplicación de redes neuronales long short term memory (LSTM), con el fin de calcular la fracción volumétrica de un flujo bifásico compuesto por aceite y agua en tubería horizontal. De tal modo que se entrenen diferentes configuraciones de redes neuronales, los resultados obtenidos en este proceso son comparados entre sí, de tal manera que permita establecer la precisión del modelo predictivo creado al implementar inteligencia artificial para optimizar el transporte de fluidos.



## 1 Objetivos

### 1.1 Objetivo General

Calcular la fracción volumétrica de aceite y agua en tuberías horizontales a partir de la aplicación de la técnica de aprendizaje automático Long Short Term Memory con el fin de optimizar el transporte de fluidos.

### 1.2 Objetivos Específicos

- Seleccionar los parámetros característicos para la determinación de la fracción volumétrica de aceite y agua.
- Diseñar un modelo de LSTM con funciones de activación, entrenamiento y de aprendizaje específicas para optimizar el modelo predictivo generado a partir del desarrollo matemático ejecutado en sus capas ocultas con información recopilada en una base de datos obtenida del análisis bibliométrico.
- Desarrollar una herramienta predictiva basada en LSTM para el cálculo de las fracciones volumétricas de aceite y agua al interior de tuberías horizontales utilizando el software libre Python.

## 2. Materiales y métodos

Se llevó a cabo un análisis de información experimental obtenida a partir de una investigación previa (Ruiz, 2021), en la cual se estudiaron las características comportamentales de un fluido bifásico. En este estudio, 13 autores se enfocaron en el comportamiento de una mezcla de aceite y agua en tubería horizontal, y se presentaron 17 esquemas diferentes como se muestra en la Tabla 1.

**Tabla 1**

*Resumen Base de Datos Recopilada*

<b>Autor</b>	<b><math>D_i</math> [m]</b>	<b><math>\mu_o</math> [Pa. s]</b>	<b><math>\rho_o</math> [kg/m<sup>3</sup>]</b>	<b>No. de datos</b>
(O. M. H. Rodriguez, 2017)	0.021	0.799	886	138
(Abduvayt et al., 2004)	0.1064	0.00188	800	49
(Cai et al., 2012)	0.1	0.002	825	99
(K. García, J. García, G. Montoya, M. T. Valecillos, n.d.)	0.0445	0.1077	881	27
	0.0508	0.0288	884	106
(Al-Sarkhi et al., 2017)	0.0508	0.013	858.5	30
	0.0508	0.013	858.5	88
(Shi et al., 2017)	0.026	5.6	910	69
(Al-Wahaibi et al., 2014)	0.019	0.012	875	196
	0.0254	0.012	875	268
(Dasari et al., 2013)	0.025	0.107	889	536
(Ibarra et al., 2015)	0.032	0.0054	825	57
(Nädler & Mewes, 1997)	0.059	0.022	----	76

(Rodriguez & Oliemans, 2006)	0.0828	0.00717	831.4	43
(Qiao et al., 2012)	0.026	5	910	64
	0.0056	0.0052	818	213
(Wegmann & Rudolf von Rohr, 2006)	0.007	0.0052	818	97

*Nota.* Adaptado de Información general de la base de datos de la literatura para flujo bifásico de aceite y agua, de Carlos Ruiz, 2021, (Ruiz, 2021).

De igual manera es de resaltar que las variables en las que enfatizó cada autor en su estudio complementando la información suministrada por la Tabla 1 son, las velocidades superficiales de aceite ( $J_o$ ) y del agua ( $J_w$ ), asimismo como la sumatoria de las velocidades superficiales de aceite y agua ( $J_o + J_w$ ) y la fracción volumétrica tanto del aceite ( $\epsilon_o$ ) como del agua ( $\epsilon_w$ ) en el fluido.

## 2.1. Tratamiento de los datos

Para realizar el preprocesamiento de los datos se utilizó el software libre de Python, más específicamente la versión presente en Google Colab donde se ingresó la base de datos y utilizando librerías como Pandas y Matplotlib se realiza el análisis y visualización de los datos.

El apropiado proceso de organización y preprocesamiento de los datos brutos fue fundamental para lograr una adecuada transformación y limpieza de los datos. Dado que la información se encontraba en diferentes escalas, se realizó un proceso de normalización de la base de datos con el fin de compararlos debidamente y evitar que los valores extremos dominen en el análisis.

## 2.2. Diseño de la red neuronal

Se separó el conjunto de datos de tal manera que se pudiera obtener el vector de entrada con la información requerida y así mismo el vector de salida. En este proceso se especificó que el

vector de entrada tendría una dimensión de cinco filas como se puede observar en (1), excluyendo el dato conocido como la fracción volumétrica del aceite.

$$X = \begin{bmatrix} J_o \\ J_w \\ J_o + w \\ \mu_o \\ D \end{bmatrix} \quad (1)$$

De igual manera se realizó el proceso de organizar el vector de salida del sistema, conformado por la selección del dato de la fracción volumétrica del aceite.

$$Y = [\varepsilon_o] \quad (2)$$

Una vez que se logró la correcta vectorización de la información, se procedió a la segmentación de los datos para el entrenamiento y la prueba del sistema, utilizando una proporción del 80% y 20% de la información total, respectivamente. Estos valores se seleccionaron para mantener el equilibrio del sistema. Por lo tanto, los datos de entrenamiento consistieron en 1724 registros independientes, mientras que los datos de prueba constaron de 432 registros para evaluar el rendimiento del sistema.

Para este estudio, se propone el uso de una red neuronal recurrente (RNN) con el objetivo de desarrollar un modelo predictivo de la fracción volumétrica de aceite y agua en un entorno bifásico. La arquitectura de la RNN es capaz de procesar y predecir secuencias de datos gracias a su función de memoria y conexiones retroalimentadas, lo que la diferencia de las redes neuronales convencionales.

Las Redes de Memoria a Largo Plazo (LSTM), son una variante especial de las Redes Neuronales Recurrentes (RNN), diseñadas para lograr una mejor predicción de secuencias de datos a largo plazo. Esto se debe a la presencia de celdas de memoria que actualizan su

información interna en cada ciclo, permitiendo una mejor retención de información relevante a lo largo de la secuencia. En este tipo de modelos, se requiere que la red tenga al menos una capa oculta para lograr una buena predicción. Las neuronas en cada capa oculta utilizan el proceso de retropropagación "Backpropagation" para ajustar los pesos sinápticos de las conexiones neuronales, con el objetivo de minimizar el error de predicción en la salida de la red. Este ajuste se realiza desde la capa de salida hasta la capa de entrada. Después del ajuste de los pesos sinápticos, se utiliza una función de transferencia de tipo ReLu para mejorar la precisión en la determinación del valor de salida deseado. Matemáticamente, la función de entrada neta se puede expresar como (3).

$$I(t) = f[w_i * x_t + U_i * h_{t-1} + b_i] \quad (3)$$

Donde:

$f$ : Función relu

$w_i$ : Matriz de pesos sinápticos

$x_t$ : Entrada en el tiempo  $t$

$U_i$ : Matriz de pesos de la salida anterior

$h_{t-1}$ : Salida de neurona en el tiempo  $t - 1$

$b_i$ : Vectores de sesgo bias

La selección adecuada de las funciones de activación fue un paso crucial en el desarrollo de la investigación. Teniendo en cuenta la naturaleza de los datos y el enfoque del estudio, se decidió utilizar las funciones de activación Rectified Linear Unit (ReLu), Sigmoide (Sigmoid) y Tangente Hiperbólica (Tanh). Dado que el intervalo de acción de las dos primeras funciones se encuentra en  $[0,1]$  y  $[-1,1]$  para la tercera función, se cumplen con los requisitos necesarios para el análisis de los datos proporcionados. Estas funciones están definidas por:

$$ReLU(I_o) = \begin{cases} 0 & ; I_o \leq 0 \\ I_o & ; I_o > 0 \end{cases} \quad (4)$$

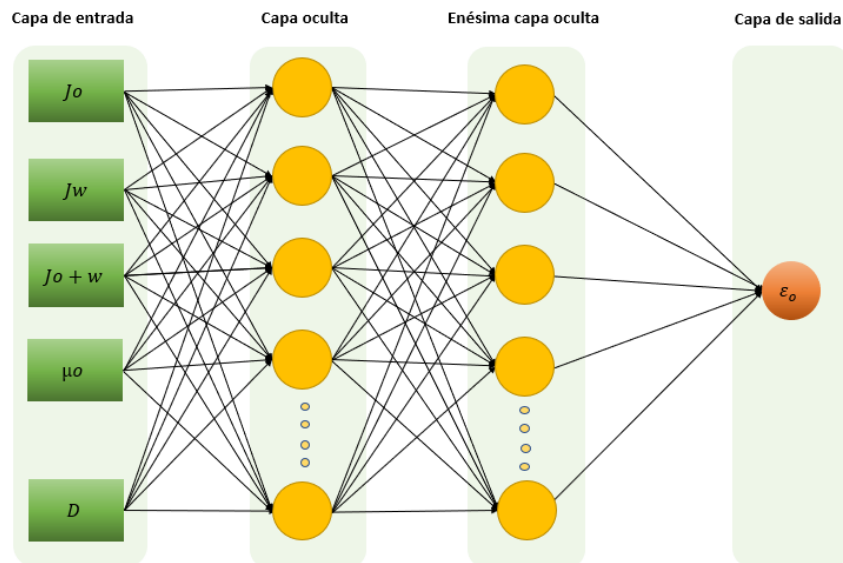
$$Sigmoid(I_o) = \frac{1}{1 + e^{-I_o}} \quad (5)$$

$$Tanh(I_o) = \frac{1 - e^{-I_o}}{1 + e^{-I_o}} \quad (6)$$

Se describe la función de salida ( $I_o$ ), que representa el resultado obtenido después de que la neurona ha completado su proceso y será utilizada como entrada en la siguiente neurona de la siguiente capa oculta. El modelo propuesto se implementó utilizando las librerías de Python especializadas en machine learning y procesamiento de datos, como Sklearn, Tensorflow y Keras.

### Figura 1

*Representación esquemática de la estructura de red neuronal artificial utilizada*



*Nota.* Tomada de elaboración propia.

Durante el proceso de entrenamiento del modelo, pueden surgir errores que están vinculados con diferentes factores, como la cantidad de neuronas, los optimizadores, las

funciones de activación y la cantidad de datos. Para determinar la precisión del modelo, se utiliza el error cuadrático medio (MSE) como el principal parámetro de evaluación, el cual se calcula de la siguiente manera.

$$MSE = \frac{1}{n} \sum_{m=1}^n (Y_{(Exp,m)} - Y_{(Pred,m)})^2 \quad (7)$$

Donde:

$n$ : Número total de datos de entrada

$Y_{(Exp)}$  = Valor experimental de la salida

$Y_{(Pred)}$  = Valor de salida obtenido por la RNA

Además de utilizar el error cuadrático medio (MSE) para evaluar la precisión del modelo, también se propone emplear el error absoluto medio (MAE) y el error absoluto medio porcentual (MAPE). La inclusión de estos parámetros permitirá validar la información, comparar los resultados y determinar las configuraciones que presentan mayor precisión.

$$MAE = \frac{1}{n} \sum_{m=1}^n |(Y_{(Exp,m)} - Y_{(Pred,m)})| \quad (8)$$

$$MAPE = \frac{100\%}{n} \sum_{m=1}^n \left| \frac{(Y_{(Exp,m)} - Y_{(Pred,m)})}{Y_{(Exp,m)}} \right| \quad (9)$$

En este estudio se plantea un análisis de experimentos que permite obtener diferentes configuraciones y grupos de prueba con el fin de comparar los resultados entre ellos, teniendo en cuenta como parámetros a variar los presentados en Tabla 2.

**Tabla 2***Parámetros del diseño de experimentos (DoE)*

Neurona:	Capa oculta	Learning Épocas	Learning Rate	Batch size	Transferencia	Optimizador	Activación salida	Métricas error
150	2	50	0.1	3	Relu	ADAM	Relu	MSE
100		20	0.2	1		RMSprop	Sigmoid	MAE
50		10					Tanh	MAPE
25								

*Nota.* Tomado de elaboración propia.

Se plantea la realización de pruebas iterando con cada una de las posibles configuraciones que involucren los parámetros presentados en Tabla 2, con el fin de poder comparar el comportamiento del modelo, teniendo en cuenta factores como los tiempos de cómputo de cada prueba, y de igual manera los resultados presentados por las métricas de error MSE, MAE y MAPE según corresponda y de acuerdo a estos resultados poder identificar las mejores configuraciones del modelo.

### 3. Resultados

Teniendo en cuenta los parámetros presentados en Tabla 2, donde se planteaba la metodología a seguir para obtener el rendimiento de modelo y asimismo sus mejores configuraciones, se presentan los resultados tomando como parámetro de segmentación el número de neuronas usadas en las capas ocultas de cada modelo. Realizando así un proceso de 216 pruebas donde se evidenció el comportamiento del sistema al realizar la respectiva comparación de los datos entregados por el modelo con los datos de validación antes separados para este fin.



En la Tabla 3

*Rendimiento del modelo variando los parámetros principales, usando 150 neuronas para su funcionamiento para las 10 configuraciones con el mejor MSE.* se presentan las 10 configuraciones con mejor MSE, estructuradas en sus capas ocultas con 150 neuronas. El presente resultado muestra que teniendo este tipo de configuración se pueden alcanzar un MSE de 6.5416E-06, de igual manera un error absoluto medio (MAE) de 0.0058 y un error absoluto medio porcentual de 3.3738%.

**Tabla 3**

*Rendimiento del modelo variando los parámetros principales, usando 150 neuronas para su funcionamiento para las 10 configuraciones con el mejor MSE.*

Pr.	Neuronas	Épocas	Learning Rate	Batch size	Optimizador	Activación salida	Loss (MSE)	MAPE	MAE
4	150	50	0.1	3	ADAM	Relu	6,5416E-06	3,3738	0,0058
11	150	50	0,1	1	ADAM	Sigmoid	1,1083E-05	1,4288	0,0022
23	150	50	0,2	1	ADAM	Sigmoid	1,7733E-05	1,6007	0,0028
24	150	50	0,2	1	ADAM	Tanh	3,0537E-05	2,9303	0,0033
35	150	20	0,1	1	ADAM	Sigmoid	3,5651E-05	3,0250	0,0045
8	150	50	0,1	1	RMSprop	Sigmoid	3,5835E-05	3,2018	0,0038
20	150	50	0,2	1	RMSprop	Sigmoid	4,1199E-05	2,3200	0,0034
12	150	50	0,1	1	ADAM	Tanh	4,2630E-05	3,1312	0,0054
36	150	20	0,1	1	ADAM	Tanh	4,3870E-05	3,5740	0,0051
6	150	50	0,1	3	ADAM	Tanh	4,5378E-05	3,1567	0,0054

*Nota.* Tomado de elaboración propia.

Según se observa en la Tabla 3, las diferentes configuraciones implementadas en el sistema generan comportamientos distintivos en los resultados. Se debe tener en cuenta que uno

de los parámetros que se varió fue el número máximo de épocas del sistema, el cual se fijó en 50. Los resultados indican que no hay una gran diferencia en el valor del MSE y MAPE el cual para 50 épocas tiene un valor de 6.5416E-06 y 3,3738% respectivamente y al realizar una prueba con 20 épocas de 3.5651E-05 y 3.0250%, lo que según la métrica porcentual no llega a ser mayor al 0.4% de mayor rendimiento, pero sí se aumenta el consumo de recursos computacionales entre 4 y 5 minutos. Debido a lo anterior, se decidió reducir el número de épocas a 20 y 10 para futuras pruebas según sea necesario.

**Tabla 4**

*Resultados rendimiento del modelo teniendo en cuenta una variación completa de parámetros principales, usando 100 neuronas para su funcionamiento*

Pr.	Neuronas	Épocas	Learning Rate	Batch size	Optimizador	Activación salida	Loss (MSE)	MAPE	MAE
78	100	20	0,1	3	ADAM	Tanh	6,3856E-05	4,0851	0,0048
77	100	20	0,1	3	ADAM	Sigmoid	6,5345E-05	3,6244	0,0054
81	100	20	0,1	1	RMSprop	Tanh	6,7145E-05	3,3238	0,0045
89	100	20	0,2	3	ADAM	Sigmoid	6,9269E-05	2,9174	0,0050
83	100	20	0,1	1	ADAM	Sigmoid	8,0392E-05	4,2026	0,0058
94	100	20	0,2	1	ADAM	Relu	8,6029E-05	3,9164	0,0067
80	100	20	0,1	1	RMSprop	Sigmoid	9,1156E-05	3,6893	0,0061
82	100	20	0,1	1	ADAM	Relu	1,0608E-04	5,8229	0,0076
107	100	10	0,1	1	ADAM	Sigmoid	1,1265E-04	4,9202	0,0089
104	100	10	0,1	1	RMSprop	Sigmoid	1,2577E-04	4,5328	0,007

*Nota.* Tomado de elaboración propia.

En la Tabla 4 se puede evidenciar que el rendimiento del modelo con una arquitectura de 100 neuronas en su capa oculta y con 20 épocas presenta un MSE de  $6.3856E-05$  comparado con el  $3.5651E-05$  presentado en la Tabla 3 al implementar 20 épocas. De igual manera se puede evidenciar que el MAPE presenta un rendimiento del 4.0851% y un MAE de 0.0048, obteniendo un menor rendimiento al presentado las configuraciones en la Tabla 3. Se puede evidenciar que el uso de los optimizadores aún no presenta una diferencia marcada para inclinarse por el uso de uno de los dos.

En las Tabla 3

*Rendimiento del modelo variando los parámetros principales, usando 150 neuronas para su funcionamiento para las 10 configuraciones con el mejor MSE.* se realizó el mismo procedimiento, mientras se variaba como dato principal el rendimiento del sistema y el número de neuronas según corresponda, obteniendo los respectivos resultados de las métricas de comparación requeridos.

En la Tabla 5 se realiza el proceso de variación de parámetros, teniendo como principal que se implementaron 50 neuronas en su arquitectura de la capa oculta. El sistema presenta un comportamiento donde en comparación con la Tabla 4, el rendimiento del MSE y MAPE presentan un aumento, llegando en este caso a ser de  $5.5920E-05$  y 3.3392% respectivamente, mientras que el MAE se mantiene estable. Esta comparativa aún sigue sin superar el rendimiento presentado en Tabla 3. Por otra parte se evidencia que el modelo tiene mejor rendimiento en esta configuración al usar un Batch size de 1 al igual que un learning rate de 0,1 realizando 20 épocas en el proceso e implementando funciones de activación como lo son sigmoide y tangente hiperbólica.

**Tabla 5**

*Rendimiento del modelo variando los parámetros principales, usando 50 neuronas para su funcionamiento para las 10 configuraciones con el mejor MSE*

Pr.	Neuronas	Épocas	Learning Rate	Batch size	Optimizador	Activación salida	Loss (MSE)	MAPE	MAE
131	50	20	0,1	1	ADAM	Sigmoid	5,5920E-05	3,3392	0,0048
128	50	20	0,1	1	RMSprop	Sigmoid	6,9462E-05	3,6845	0,0047
132	50	20	0,1	1	ADAM	Tanh	7,8628E-05	6,2786	0,0062
129	50	20	0,1	1	RMSprop	Tanh	9,3823E-05	5,3622	0,0063
137	50	20	0,2	3	ADAM	Sigmoid	9,8305E-05	3,5328	0,0059
125	50	20	0,1	3	ADAM	Sigmoid	1,0061E-04	4,1487	0,0074
142	50	20	0,2	1	ADAM	Relu	1,0675E-04	4,0291	0,0074
156	50	10	0,1	1	ADAM	Tanh	1,1051E-04	3,9565	0,0074
153	50	10	0,1	1	RMSprop	Tanh	1,2646E-04	3,4953	0,0078
127	50	20	0,1	1	RMSprop	Relu	1,2680E-04	5,1179	0,0077

*Nota.* Tomado de elaboración propia.

Los resultados presentados en la Tabla 6 donde el parámetro de agrupación presenta una estructura de 25 neuronas en su capa oculta, se puede evidenciar la disminución del rendimiento del MSE y MAE con respecto los resultados anteriores, presentando un error de 8.2504E-05 y 0,0063 respectivamente. Respecto al análisis del MAPE se puede evidenciar una mejora con

respecto a la Tabla 5 dado que este presenta 3,1885% frente a los 3.3392% del resultado anterior, presentando una mejora del 0,15%.

**Tabla 6**

*Rendimiento del modelo variando los parámetros principales, usando 25 neuronas para su funcionamiento para las 10 configuraciones con el mejor MSE.*

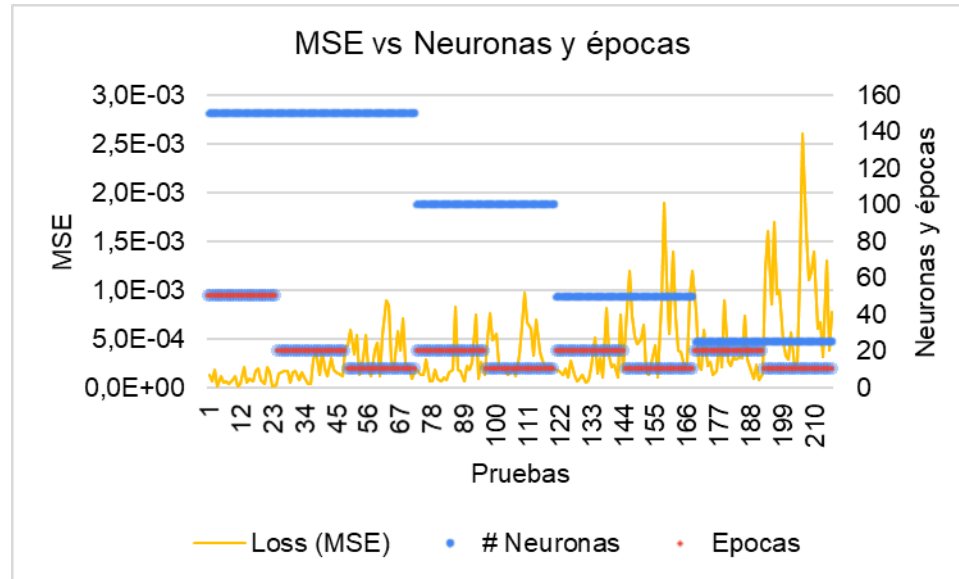
Pr.	Neuronas	Épocas	Learning Rate	Batch size	Optimizador	Activación salida	Loss (MSE)	MAPE	MAE
191	25	20	0,2	1	ADAM	Sigmoid	8,2504E-05	3,1885	0,0063
189	25	20	0,2	1	RMSprop	Tanh	8,5745E-05	2,651	0,006
175	25	20	0,1	1	RMSprop	Relu	1,3179E-04	5,431	0,0078
192	25	20	0,2	1	ADAM	Tanh	1,4185E-04	4,5988	0,0083
188	25	20	0,2	1	RMSprop	Sigmoid	1,6778E-04	4,3604	0,0095
176	25	20	0,1	1	RMSprop	Sigmoid	1,7029E-04	4,7803	0,0095
171	25	20	0,1	3	RMSprop	Tanh	1,8989E-04	5,8035	0,0098
203	25	10	0,1	1	ADAM	Sigmoid	2,0533E-04	6,3683	0,0096
178	25	20	0,1	1	ADAM	Relu	2,0710E-04	7,7444	0,0101
173	25	20	0,1	3	ADAM	Sigmoid	2,2334E-04	6,0608	0,0092

*Nota.* Tomada de elaboración propia.

En la Figura 2 se presenta el comportamiento del error cuadrático medio y la relación que presenta con los parámetros de variación número de neuronas y número de épocas del sistema. Se puede observar que durante el desarrollo de las 216 pruebas no se presentó una tendencia marcada del comportamiento del sistema, por lo que se requiere hacer un análisis más detallado del funcionamiento del modelo.

**Figura 2**

*Comportamiento de MSE relacionado con el número de neuronas y épocas del sistema*

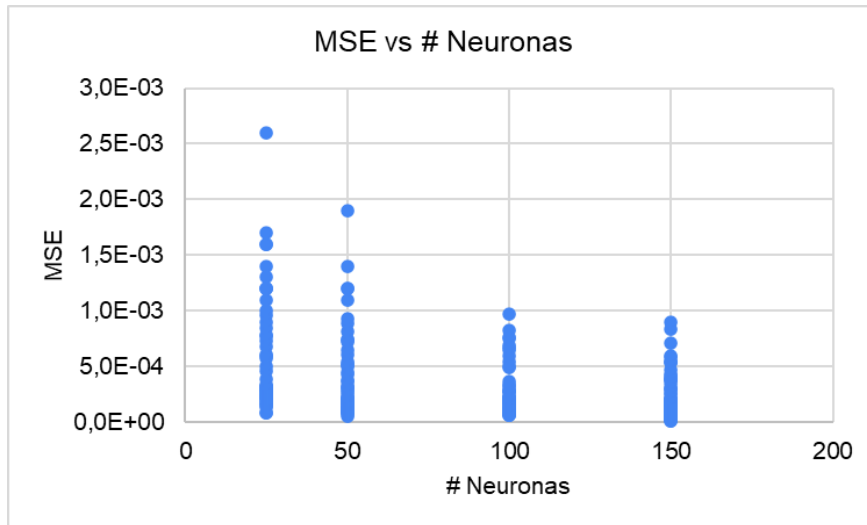


*Nota.* Tomado de elaboración propia.

Con el fin de realizar un análisis más detallado de los resultados de rendimiento del sistema se planteó en Figura 3 realizar una organización de menor a mayor de los valores del MSE y compararlos en el número de neuronas para poder revisar la correlación entre ellos y evidenciar la influencia de este parámetro en el rendimiento del modelo.

**Figura 3**

*Correlación entre el número de neuronas en la capa oculta y el MSE del sistema*



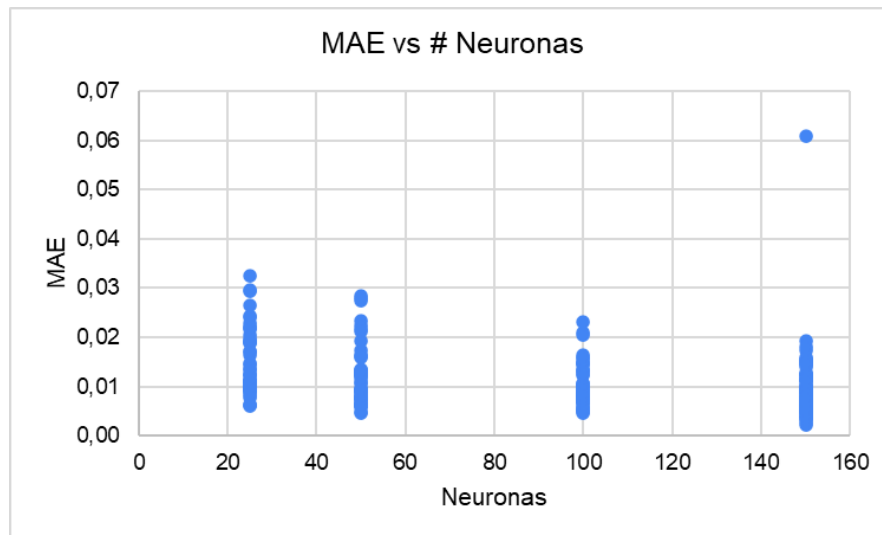
*Nota.* Tomado de elaboración propia.

Se puede evidenciar en Figura 3 que con las configuraciones de 100 y 150 neuronas se consiguen mejores rendimientos de MSE en el sistema, que implementando variaciones con 50 o 25 neuronas respectivamente.

En la Figura 4 se presenta la correlación entre el número de neuronas implementados en el sistema y el resultado de la métrica error absoluto medio (MAE). Donde se puede evidenciar que, a diferencia de los resultados presentados por el MSE, los cuales estaban mucho más segmentados, para este caso los valores son muy similares, sin embargo, es posible divisar una pequeña diferencia entre los sistemas con 150 y 100 neuronas que lo que contienen 50 y 25 respectivamente. Con este análisis se soporta lo mencionado en la Figura 3.

#### **Figura 4**

*Variación del error absoluto medio (MAE) teniendo en cuenta el número de épocas del sistema*



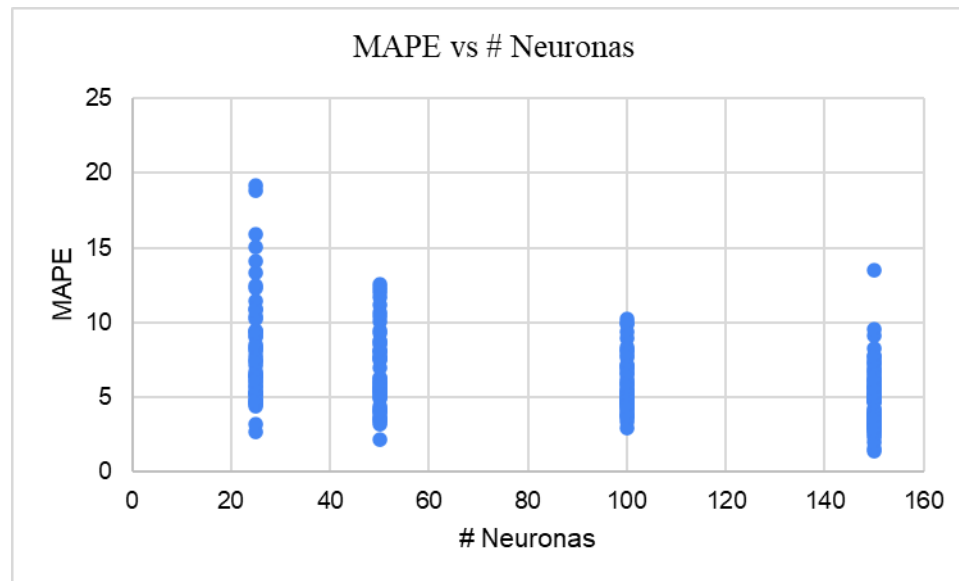
*Nota.* Tomado de elaboración propia

Con el fin de concluir la interpretación de resultados del modelo, se presentó la Figura 5, donde se compararon los resultados de MAPE y cuál es su incidencia con el número de neuronas presentes en el sistema. Dando como resultado que de igual manera como sus antecesoras, las configuraciones que presentaban 150 y 100 neuronas en su estructura interna, tienden a presentar menor error en el momento realizar el entrenamiento del modelo, de igual manera los resultados pertenecientes a las configuraciones de 50 y 25 neuronas no están mal diseñados, solo que generaron menos eficiencia que las variaciones antes mencionadas.

### **Figura 5**

*Variación del error absoluto medio porcentual (MAPE) teniendo en cuenta el número de neuronas del sistema*



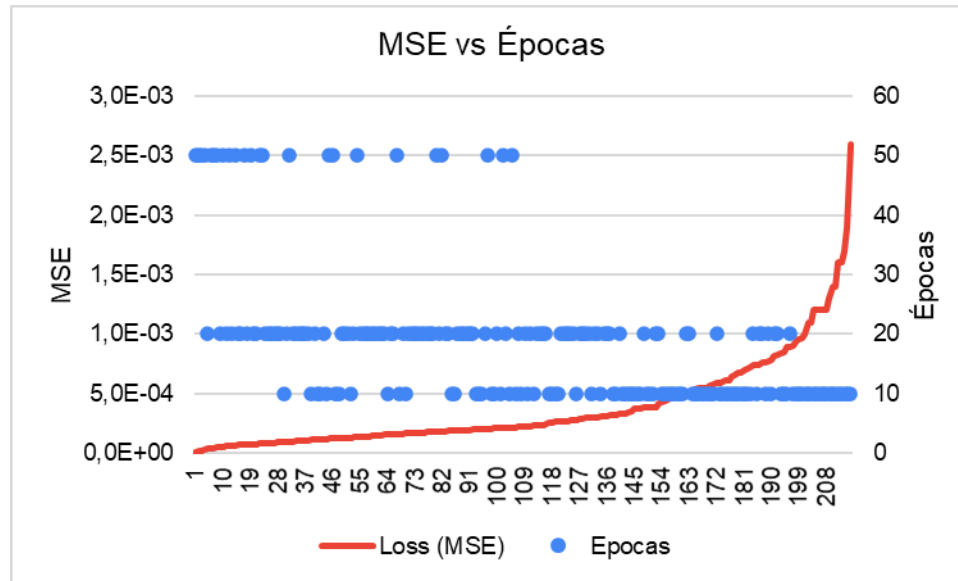


*Nota.* Tomado de elaboración propia.

Para poder estimar cuales fueron los valores óptimos de trabajo al realizar este estudio se procedió a analizar el comportamiento del número de épocas en comparación con la métrica de error MSE como se muestra en Figura 6, esto teniendo en cuenta los resultados antes obtenidos de cada uno de los análisis relacionados al número de neuronas del sistema. Para mejor interpretación de los datos obtenidos se realizó un ordenamiento de la información de menor a mayor, con esto se presenta una tendencia creciente del gráfico en cuestión.

### **Figura 6**

*Análisis de comportamiento del error cuadrático medio (MSE), contra el número de épocas*



*Nota.* Tomado de elaboración propia.

Como se puede evidenciar en la Figura 6, dado que los resultados del MSE presentan una función con un comportamiento creciente, de igual manera y retomando el análisis realizado en Tabla 3, se comprueba que al implementar 50 épocas en el sistema, no genera gran impacto en los resultados, de igual manera, dada la interpretación se evidencia que cuando se implementan 10 épocas en el sistema, se presentan bajos rendimientos y errores altos, por lo que se tiene que para el estudio, el parámetro de 20 épocas es el que mejor cumple con un buen rendimiento y optimización de recursos de cómputo.

Después de haber realizado el análisis de resultados, y con el fin de obtener la variación de parámetros más óptima se planteó realizar la selección de las 10 mejores configuraciones en, teniendo en cuenta que por motivo de eficiencia y como se mencionó en Tabla 3, no se tomarán en cuenta las pruebas con 50 épocas en su desarrollo.

**Tabla 7**

*Resultados generales de rendimiento teniendo como parámetro de referencia el MSE*

Pr.	Neuronas	Épocas	Learning Rate	Batch size	Optimizador	Activación salida	Loss (MSE)	MAPE	MAE
35	150	20	0,1	1	ADAM	Sigmoid	3,5651E-05	3,0250	0,0045
36	150	20	0,1	1	ADAM	Tanh	4,3870E-05	3,5740	0,0051
131	50	20	0,1	1	ADAM	Sigmoid	5,5920E-05	3,3392	0,0048
29	150	20	0,1	3	ADAM	Sigmoid	5,8223E-05	3,3211	0,0050
78	100	20	0,1	3	ADAM	Tanh	6,3856E-05	4,0851	0,0048
77	100	20	0,1	3	ADAM	Sigmoid	6,5345E-05	3,6244	0,0054
81	100	20	0,1	1	RMSprop	Tanh	6,7145E-05	3,3238	0,0045
89	100	20	0,2	3	ADAM	Sigmoid	6,9269E-05	2,9174	0,0050
128	50	20	0,1	1	RMSprop	Sigmoid	6,9462E-05	3,6845	0,0047
132	50	20	0,1	1	ADAM	Tanh	7,8628E-05	6,2786	0,0062

*Nota.* Tomado de elaboración propia

Como se puede observar, en la Tabla 7, se presentan combinaciones de redes neuronales con estructura interna de 150, 100 y 50 neuronas, de igual manera las épocas del sistema se mantuvieron constantes en 20, presentando mayoría en el uso del optimizador ADAM y como función de activación se presenta la sigmoide. Teniendo como la mejor opción, con el fin de obtener mejor rendimiento el procedimiento realizado en la prueba 35, la cual contaba con los siguientes parámetros Tabla 8.

**Tabla 8**

*Selección de parámetros del sistema para obtener el mejor rendimiento del estudio*

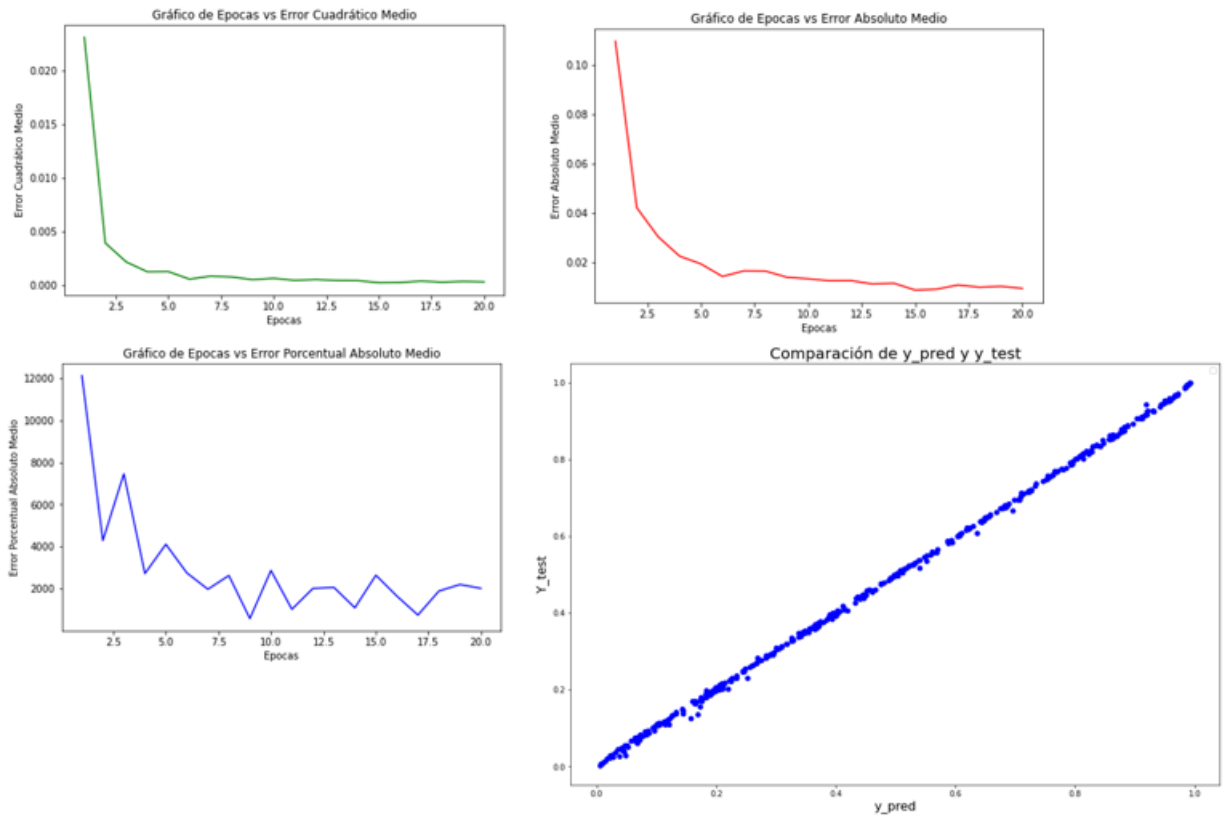
<b>Pr.</b>	<b>Neuronas</b>	<b>Épocas</b>	<b>Learning Rate</b>	<b>Batch size</b>	<b>Optimizador</b>	<b>Activación salida</b>	<b>Loss (MSE)</b>	<b>MAPE</b>	<b>MAE</b>
35	150	20	0,1	1	ADAM	Sigmoid	3,5651E-05	3,0250	0,0045

*Nota.* Tomado de elaboración propia.

La Figura 7 presenta las respectivas gráficas de rendimiento del modelo seleccionado, se presentan 4 gráficas, comparando el número de épocas con el MSE, de igual manera el error absoluto medio (MAE), el error porcentual absoluto medio (MAPE) y se presenta la comparación de los resultados entre los datos predichos por el modelo y los datos dispuestos para test.

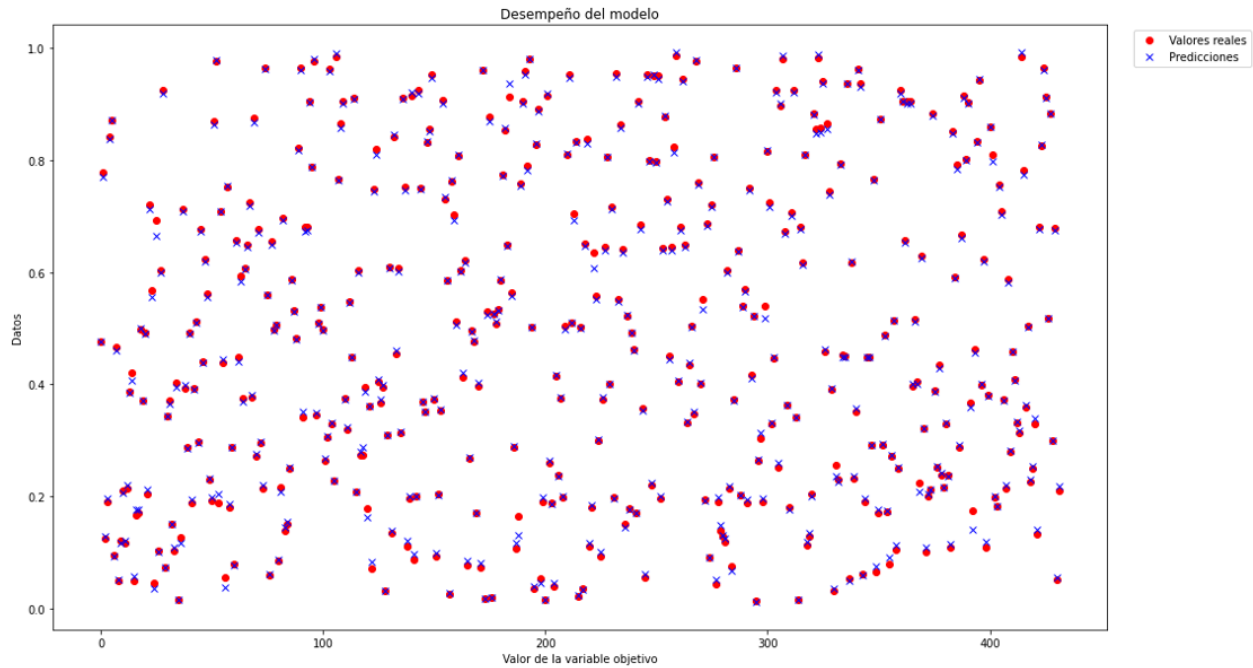
Figura 7

*Gráficas de rendimiento de la configuración seleccionada*



*Nota.* Tomado de elaboración propia.

En complemento de la información anterior se presenta la Figura 8, donde se muestra el comportamiento de los datos predichos y los 432 datos dispuestos para validar la información, esto con el fin de visualizar el funcionamiento del sistema y su validación.

**Figura 8***Desempeño del modelo seleccionado*

*Nota.* Tomado de elaboración propia

#### 4. Conclusiones

Se definió el mejor modelo de RNA que permitió calcular la fracción volumétrica de aceite y agua en tubería horizontal a través del procesamiento y predicción de series de datos. El modelo seleccionado presentó un vector de entrada que le incluía velocidades superficiales de cada fluido, diámetro interno de la tubería, velocidad de la mezcla y viscosidad. Según las pruebas de rendimiento, el modelo presenta un error absoluto medio porcentual de 3.02% y un error absoluto medio de 0.0045.

Se implementó una red neuronal Long Short Term Memory (LSTM), donde se evidenció el alcance y versatilidad de los modelos LSTM en el área de la predicción de series temporales.

El modelo que por su buen desempeño fue elegido como el mejor de este estudio, consta de 150 neuronas con función de transferencia ReLu, con, 2 capas ocultas, optimizador ADAM y una capa de salida para la predicción de la fracción volumétrica del aceite.

Con la implementación de un buen modelo de redes recurrentes LSTM en el software libre Python, pueden predecir la fracción volumétrica de un fluido bifásico, de igual manera es fundamental realizar un preprocesamiento correcto de los datos para así optimizar los resultados y la eficiencia del sistema.

## 5. Referencias

- Abdelhafidh, M., Fourati, M., Fourati, L. C., & Abidi, A. (2018). Remote water pipeline monitoring system IoT-based architecture for new industrial era 4.0. *Proceedings of IEEE/ACS International Conference on Computer Systems and Applications, AICCSA, 2017-October*, 1184–1191. <https://doi.org/10.1109/AICCSA.2017.158>
- Abduvayt, P., Manabe, R., Watanabe, T., & Arihara, N. (2004). Analisis of oil-water flow tests in horizontal, hilly-terrain, and vertical pipes. *Proceedings - SPE Annual Technical Conference and Exhibition*, 1335–1347. <https://doi.org/10.2523/90096-ms>
- Al-Sarkhi, A., Pereyra, E., Mantilla, I., & Avila, C. (2017). Dimensionless oil-water stratified to non-stratified flow pattern transition. *Journal of Petroleum Science and Engineering*, 151(September 2016), 284–291. <https://doi.org/10.1016/j.petrol.2017.01.016>
- Al-Wahaibi, T., Al-Wahaibi, Y., Al-Ajmi, A., Al-Hajri, R., Yusuf, N., Olawale, A. S., & Mohammed, I. A. (2014). Experimental investigation on flow patterns and pressure gradient through two pipe diameters in horizontal oil-water flows. *Journal of Petroleum Science and Engineering*, 122, 266–273. <https://doi.org/10.1016/j.petrol.2014.07.019>
- Bikmukhametov, T., & Jäschke, J. (2020). First Principles and Machine Learning Virtual Flow Metering: A Literature Review. In *Journal of Petroleum Science and Engineering* (Vol. 184). Elsevier B.V. <https://doi.org/10.1016/j.petrol.2019.106487>
- Biswas, D. (2019). Adapting shallow and deep learning algorithms to examine production performance – Data analytics and forecasting. *Society of Petroleum Engineers - SPE/IATMI Asia Pacific Oil and Gas Conference and Exhibition 2019, APOG 2019*. <https://doi.org/10.2118/196404-ms>



- Cai, J., Li, C., Tang, X., Ayello, F., Richter, S., & Nestic, S. (2012). Experimental study of water wetting in oil-water two phase flow-Horizontal flow of model oil. *Chemical Engineering Science*, 73, 334–344. <https://doi.org/10.1016/j.ces.2012.01.014>
- Cao, Q., Banerjee, R., Gupta, S., Li, J., Zhou, W., & Jeyachandra, B. (2016). Data driven production forecasting using machine learning. *Society of Petroleum Engineers - SPE Argentina Exploration and Production of Unconventional Resources Symposium*. <https://doi.org/10.2118/180984-ms>
- Chaki, S., Verma, A. K., Routray, A., Mohanty, W. K., & Jenamani, M. (2014). Well tops guided prediction of reservoir properties using modular neural network concept: A case study from western onshore, India. *Journal of Petroleum Science and Engineering*, 123, 155–163. <https://doi.org/10.1016/j.petrol.2014.06.019>
- Consuegra, F. R., Enrique, M., & Simancas, C. (2017). *Sistemas De Recolección De Crudo*. 15, 87–99.
- Dasari, A., Desamala, A. B., Dasmahapatra, A. K., & Mandal, T. K. (2013). Experimental studies and probabilistic neural network prediction on flow pattern of viscous oil-water flow through a circular horizontal pipe. *Industrial and Engineering Chemistry Research*, 52(23), 7975–7985. <https://doi.org/10.1021/ie301430m>
- Ganeshmoorthy, V., Muthukannan, M., & Thirugnanasambandam, M. (2019). Comparison of Linear Regression and ANN of Fish Oil Biodiesel Properties Prediction. *Journal of Physics: Conference Series*, 1276(1). <https://doi.org/10.1088/1742-6596/1276/1/012074>
- Ibarra, R., Zadrazil, I., Markides, C. N., & Matar, O. K. (2015). Towards a universal dimensionless map of flow regime transitions in horizontal liquid-liquid flows. *11th International Conference on Heat Transfer, Fluid Mechanics and Thermodynamics*, July.

[https://www.researchgate.net/publication/283498693\\_Towards\\_a\\_Universal\\_Dimensionless\\_Map\\_of\\_Flow\\_Regime\\_Transitions\\_in\\_Horizontal\\_Liquid-Liquid\\_Flows](https://www.researchgate.net/publication/283498693_Towards_a_Universal_Dimensionless_Map_of_Flow_Regime_Transitions_in_Horizontal_Liquid-Liquid_Flows)

Ibrahim, A., Hewakandamby, B., Yang, Z., & Azzopardi, B. (2018). Effect of liquid viscosity on two-phase flow development in a vertical large diameter pipe. *American Society of Mechanical Engineers, Fluids Engineering Division (Publication) FEDSM*, 3, 1–10.

<https://doi.org/10.1115/FEDSM2018-83466>

K. García, J. García, G. Montoya, M. T. Valecillos, C. R. (n.d.). Determinación de altura de fase y Hold Up para flujo bifásico líquido-líquido en tuberías horizontales por medio de procesamiento de imágenes. *VIII Congreso Anual de Ingeniería ASME, Caracas, Venezuela*, 1–10.

Li, H., Misra, S., & He, J. (2020). Neural network modeling of in situ fluid-filled pore size distributions in subsurface shale reservoirs under data constraints. *Neural Computing and Applications*, 32(8), 3873–3885. <https://doi.org/10.1007/s00521-019-04124-w>

Luo, G., Tian, Y., Bychina, M., & Ehlig-Economides, C. (2018). Production optimization using machine learning in bakken shale. *SPE/AAPG/SEG Unconventional Resources Technology Conference 2018, URTC 2018, 2011*. <https://doi.org/10.15530/urtec-2018-2902505>

Madasu, S., & Rangarajan, K. P. (2019). Deep recurrent neural network DrNN model for real-time step-down analysis. *Society of Petroleum Engineers - SPE Reservoir Characterisation and Simulation Conference and Exhibition 2019, RCSC 2019*.

<https://doi.org/10.2118/196621-ms>

Malayeri, M. R., Müller-Steinhagen, H., & Smith, J. M. (2003). Neural network analysis of void fraction in air/water two-phase flows at elevated temperatures. *Chemical Engineering and Processing: Process Intensification*, 42(8–9), 587–597. <https://doi.org/10.1016/S0255->

2701(02)00208-8

Nädler, M., & Mewes, D. (1997). Flow induced emulsification in the flow of two immiscible liquids in horizontal pipes. *International Journal of Multiphase Flow*, 23(1), 55–68.

[https://doi.org/10.1016/s0301-9322\(96\)00055-9](https://doi.org/10.1016/s0301-9322(96)00055-9)

O. M. H. Rodriguez, R. V. A. O. (2017). Experimental study on oil-water flow in horizontal and slightly inclined pipes. *International Journal of Multiphase Flow*, 32(July), 950–965.

<https://doi.org/10.1016/j.ijmultiphaseflow.2017.07.005>

Osarogiagbon, A. U., Khan, F., Venkatesan, R., & Gillard, P. (2021). Review and analysis of supervised machine learning algorithms for hazardous events in drilling operations. *Process Safety and Environmental Protection*, 147, 367–384.

<https://doi.org/10.1016/j.psep.2020.09.038>

Osundare, O. S., Falcone, G., Lao, L., & Elliott, A. (2020). Liquid-liquid flow pattern prediction using relevant dimensionless parameter groups. *Energies*, 13(17).

<https://doi.org/10.3390/en13174355>

Özgür, C., & Tosun, E. (2017). Prediction of density and kinematic viscosity of biodiesel by artificial neural networks. *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, 39(10), 985–991. <https://doi.org/10.1080/15567036.2017.1280563>

Qiao, Z., Wang, Z., Zhang, C., Yuan, S., Zhu, Y., & Wang, J. (2012). PVAm–PIP/PS composite membrane with high performance for CO<sub>2</sub>/N<sub>2</sub> separation. *AIChE Journal*, 59(4), 215–228.

<https://doi.org/10.1002/aic>

Rodriguez, O. M. H., & Oliemans, R. V. A. (2006). Experimental study on oil-water flow in horizontal and slightly inclined pipes. *International Journal of Multiphase Flow*, 32(3),

323–343. <https://doi.org/10.1016/j.ijmultiphaseflow.2005.11.001>

- Ruiz, C. (2021). *Caracterización hidrodinámica de flujos multifase empleando técnicas de inteligencia artificial*. Universidad Industrial de Santander.
- Sagheer, A., & Kotb, M. (2019). Time series forecasting of petroleum production using deep LSTM recurrent networks. *Neurocomputing*, 323, 203–213.  
<https://doi.org/10.1016/j.neucom.2018.09.082>
- Sai, K. R., Nayak, P. J., Kumar, K. V. A., & Dutta, A. D. (2020). Oil Spill Management System Based on Internet of Things. *Proceedings of 2020 IEEE-HYDCON International Conference on Engineering in the 4th Industrial Revolution, HYDCON 2020*.  
<https://doi.org/10.1109/HYDCON48903.2020.9242823>
- Salgado, C. M., Dam, R. S. F., Salgado, W. L., Santos, M. C., & Schirru, R. (2021). Development of a deep rectifier neural network for fluid volume fraction prediction in multiphase flows by gamma-ray densitometry. *Radiation Physics and Chemistry*, 189(July).  
<https://doi.org/10.1016/j.radphyschem.2021.109708>
- Sattari, M. A., Roshani, G. H., & Hanus, R. (2020). Improving the structure of two-phase flow meter using feature extraction and GMDH neural network. *Radiation Physics and Chemistry*, 171(December 2019), 108725.  
<https://doi.org/10.1016/j.radphyschem.2020.108725>
- Shi, J., Lao, L., & Yeung, H. (2017). Water-lubricated transport of high-viscosity oil in horizontal pipes: The water holdup and pressure gradient. *International Journal of Multiphase Flow*, 96, 70–85. <https://doi.org/10.1016/j.ijmultiphaseflow.2017.07.005>
- Slaný, V., Lučanský, A., Koudelka, P., Mareček, J., Krčálová, E., & Martínek, R. (2020). An integrated iot architecture for smart metering using next generation sensor for water management based on lorawan technology: A pilot study. *Sensors (Switzerland)*, 20(17), 1–

23. <https://doi.org/10.3390/s20174712>

Temirchev, P., Simonov, M., Kostoev, R., Burnaev, E., Oseledets, I., Akhmetov, A., Margarit, A., Sitnikov, A., & Koroteev, D. (2020). Deep neural networks predicting oil movement in a development unit. *Journal of Petroleum Science and Engineering*, 184(August 2019), 106513. <https://doi.org/10.1016/j.petrol.2019.106513>

Varotsis, N. (1997). Prediction of gas phase volumetric variations in diphasic reservoir oils. *Journal of Canadian Petroleum Technology*, 36(8), 56–60. <https://doi.org/10.2118/97-08-05>

Vuo, U. D. E., Vuo, U. D. E., & Vuo, U. D. E. (2021). *&11 /670 IHxudo IHwzrun \$\$solhg Iru 7Khupdo ,Qiuduhg 8Qghujurxqg :Dwhu /Hdndjh*. 219–224.

Wegmann, A., & Rudolf von Rohr, P. (2006). Two phase liquid-liquid flows in pipes of small diameters. *International Journal of Multiphase Flow*, 32(8), 1017–1028. <https://doi.org/10.1016/j.ijmultiphaseflow.2006.04.001>

Yoon, A., & Kim, S. J. (2021). A deep-learning approach for predicting oscillating motion of liquid slugs in a closed-loop pulsating heat pipe. *International Journal of Heat and Mass Transfer*, 181, 121860. <https://doi.org/10.1016/j.ijheatmasstransfer.2021.121860>

Zhong, Z., Carr, T. R., Wu, X., & Wang, G. (2019). Application of a convolutional neural network in permeability prediction: A case study in the Jacksonburg-Stringtown oil field, West Virginia, USA. *Geophysics*, 84(6), B363–B373. <https://doi.org/10.1190/geo2018-0588.1>