



**UNIVERSIDAD CÉSAR VALLEJO**

FACULTAD DE INGENIERÍA Y ARQUITECTURA

ESCUELA PROFESIONAL DE INGENIERÍA DE  
SISTEMAS

“Modelo basado en redes neuronales para la predicción de  
precios de inmuebles Piura - 2021”

TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE  
INGENIERO DE SISTEMAS

**AUTOR:**

García Palacios, Carlo César (ORCID 0000-0003-2113-0579)

**ASESOR:**

Mg. Winner Agurto Marchán (ORCID 0000-0002-0396-9349)

**LÍNEA DE INVESTIGACIÓN:**

Sistemas de Información y Comunicaciones

PIURA – PERÚ

2021

## DEDICATORIA

Para la mujer que me dió la vida y confió en mi todo este tiempo, por acompañarme a lo largo de todo mi aprendizaje y en los momentos más difíciles, **gracias mamá Nelly Milagros Palacios Ruesta** y también a mi abuela **Nelly Gabriela**.

## **AGRADECIMIENTO**

### **A Dios Todopoderoso**

A mi familia por confiar en mí, los compañeros por el apoyo y por el esfuerzo en esta etapa, a el Ing. Winner Agurto por la constante asesoría brindada este año y por los conocimientos compartidos en todo este tiempo.

## ÍNDICE DE CONTENIDO

I.	INTRODUCCIÓN .....	1
II.	MARCO TEÓRICO .....	3
III.	METODOLOGÍA.....	12
	3.1 Tipo y Diseño de investigación .....	12
	3.2 Variables y Operacionalización.....	12
	3.3 Población, muestra y muestreo .....	13
	3.4 Técnicas e instrumentos de recolección de datos .....	13
	3.5 Procedimientos .....	15
	3.6 Métodos de análisis de datos .....	15
	3.7 Aspectos éticos.....	16
IV.	RESULTADOS .....	17
V.	DISCUSIÓN.....	25
VI.	CONCLUSIONES.....	28
VII.	RECOMENDACIONES .....	29
	REFERENCIAS.....	30
	ANEXOS .....	

## ÍNDICE DE TABLAS

Tabla N° 1: Técnicas e instrumentos .....	14
Tabla N° 2: Valores obtenidos parámetros de entrada en Piura .....	17
Tabla N° 3: Valores obtenidos parámetros de entrada pág. alternas .....	18
Tabla N° 4: Resumen de red neuronal construida y entrenada N°1 .....	19
Tabla N° 4: Resumen de historial de entrenamiento .....	19
Tabla N° 5: Resultado Simulación N°1 .....	20
Tabla N° 6: medición del error del modelo N°1 .....	21
Tabla N° 7: Resumen de red neuronal construida y entrenada N°2 .....	21
Tabla N° 8: Resultado de Simulación N°2 .....	22
Tabla N° 9: medición del error del modelo N°2.....	22
Tabla N° 10: Resumen de red neuronal construida y entrenada N°3 .....	23
Tabla N° 11: Resultado de Simulación N°3 .....	24
Tabla N° 12: medición del error del modelo N°3.....	24
Tabla N° 13: Resultados de predicción del modelo .....	24

## ÍNDICE DE FIGURAS

Figura 1: Ejemplo de una red neuronal completamente conectada .....	6
Figura 3: Diagrama de Dispersión de Inmuebles .....	18
Figura 4: Gráfico rendimiento del modelo durante el entrenamiento N°1 .....	20
Figura 5: Gráfico rendimiento del modelo durante el entrenamiento N°2 .....	22
Figura 6: Gráfico rendimiento del modelo durante el entrenamiento N°3 .....	23

## RESUMEN

Esta investigación tuvo como objetivo determinar la efectividad de un modelo de redes neuronales en la predicción de precios de inmuebles en Piura 2021. El enfoque de esta investigación fue cuantitativo, porque el modelo se analizó mediante estadística descriptiva y de regresión, por lo tanto, el tipo de investigación fue aplicada, el diseño de investigación fue pre experimental y de nivel explicativo. Entre los resultados se obtuvo que el modelo elegido que se implementó, fue el modelo secuencial debido a que se basa en función a varias entradas y única salida. Se obtuvo un set de entrenamiento procesado de 21000 inmuebles entre ventas realizadas y puestos en venta. Se obtuvo también que la red neuronal tuvo 3 simulaciones, donde arrojó un score de varianza de 0,8 en la tercera simulación y una efectividad de 25% como resultado mínimo, utilizando 7 variables de entrada para el modelo y para su posterior validación. En conclusión, se determinó que el modelo de RNA es efectivo no solo por las correctas configuraciones realizadas tras cada simulación, sino por el uso de métricas de sklearn para regresiones y la selección más óptima de modelo, que permitieron evaluarlo en base a su precisión durante el entrenamiento.

**Palabras claves:** Redes neuronales, modelos de predicción, Inteligencia artificial.

## ABSTRACT

The objective of this investigation was to determine the effectiveness of a neural network model in predicting real estate prices in Piura 2021. The focus of this research is quantitative, because the model was analyzed using descriptive and regression statistics; therefore, the type of research is applied; the research design is pre-experimental and explanatory. Among the results, it was found that the sequential model was the one implemented because it is based on a function of several inputs and a single output. A processed training set of 21,000 properties was obtained, including sales made and listed for sale. Information was also obtained showing that the neural network had 3 simulations, where it yielded a variance score of 0.8 in the third simulation and an effectiveness of 25% as a minimum result, using seven input variables for the model and for its subsequent validation. In conclusion, it was determined that the ANN model is effective not only because of the correct configurations made after each simulation, but also because of the use of sklearn metrics for regression and the most optimal model selection, which allowed the model to be evaluated based on its accuracy during training.

**Keywords:** Neural networks, prediction models, artificial intelligence.



## I. INTRODUCCIÓN

En la actualidad, las tecnologías de información han tomado un rol importante en diversas acciones del ser humano. Antes, muchas de estas actividades eran reservadas únicamente para el hombre, la idea que no era tomada en cuenta sobre los ordenadores tuviera la capacidad de aprender cómo tomar decisiones sobre actividades ahora es sumamente importante, y es algo de lo que se debe ser consciente, debido a que estas actividades han ido incrementándose y haciéndose más complejas a lo largo del tiempo.

Con la aparición de la Inteligencia Artificial dentro de los avances tecnológicos, esta, tomó más importancia dentro de las organizaciones. (Rouhiainen, 2018) explica que la idea de la tecnología proyectada en IA haya sido y siga siendo utilizada en diversas tareas sirve para mejorar significativamente al humano y disfrutar de una mejor eficiencia en la mayoría de aspectos de la vida. La diferencia entre la máquina y el hombre es que esta, a diferencia de él, no requiere de un descanso y desarrolla análisis de grandes cantidades de información al mismo tiempo, lo cual su uso en ciertos ámbitos se hace indispensable.

En los últimos años, Piura se ha convertido en una ciudad altamente atractiva para el sector inmobiliario, el 60% de los créditos cedidos por el Estado en el ámbito de la materia hipotecaria se localizan en el norte del Perú, siendo Piura la principal en adquisición de propiedades (Piura News, 2015). El resultado de una pandemia por el covid-19 ha hecho que los clientes especulen precios más baratos en los bienes que considera comprar y las estimaciones de inmuebles muy variada (García, y otros, 2020).

La demanda registrada por el interés de las familias de tener como hogar una propiedad moderna y con precios dentro del presupuesto, es una tarea de los agentes inmobiliarios (Guerrero, 2018). El cliente es exigente y analiza su futura compra en cuanto a características como la antigüedad del inmueble y su infraestructura para saber si su inversión se mantendrá con el pasar de los años.

Así pues, como dada la gran cantidad de datos que se generan, cabe la posibilidad de obtener un valor a partir de ellos, siendo usada la Inteligencia

Artificial, en específico, las redes neuronales para la predicción de precios de inmuebles, y en la duda de conocer cuáles serían los límites de un modelo de red neuronal. (Villada, y otros, 2016) argumentan que a diferencia de los equipos que usan algoritmos simples, las redes neuronales funcionan simulando el cerebro humano, procesando información, generalizando situaciones que no se incluyen en el entrenamiento de la misma, naciendo la pregunta ¿Cuál es la efectividad del modelo de redes neuronales en la predicción del precio de inmuebles en la ciudad de Piura? ¿Cuál es el modelo adecuado para la predicción de precios? ¿Cuáles son los datos que se deben obtener para los parámetros de entrada de la red neuronal? ¿Cuál es el porcentaje de error necesario para la predicción de precios de inmuebles?

Esta investigación se justifica el concepto que se quiere llevar a cabo, se debe tomar en cuenta la realidad problemática mencionada anteriormente, se plantea la creación de un modelo de predicción del valor de un inmueble como solución a que el precio de los inmuebles puesto en venta es a veces ofrecido por el vendedor agregando un valor muy distanciado de un valor accesible, por lo cual, genera un precio que no es el indicado y que va variando con las ofertas de los clientes. Se tomó como enfoque del machine learning debido a la amplia información sobre de la utilización de este proceso.

Cabe contextualizar que el objetivo de la presente investigación, es el Determinar la efectividad del modelo de redes neuronales en la predicción de un inmueble, en base a sus características. Como objetivos específicos se tiene Determinar el modelo a desarrollar para la red neuronal, Identificar los parámetros de entrada para el modelo de la predicción de los precios, Evaluar el porcentaje de error mínimo para que la red neuronal sea efectiva.

Estos objetivos permiten plantear la siguiente hipótesis general: El modelo de redes neuronales es efectivo en la predicción de precios de inmuebles. Y las específicas: El modelo elegido es el adecuado para desarrollar la red neuronal en la predicción de precios, los parámetros de entrada son los correctos para predecir el precio de un inmueble mediante la RNA, el modelo de red neuronal obtiene un porcentaje mínimo para predecir el valor de un inmueble.

## II. MARCO TEÓRICO

En el siguiente capítulo de investigación se realizó una búsqueda de distintos antecedentes nacionales e internacionales relacionados con el tema de investigación.

En el ámbito internacional se tiene a (Bozanic, 2020) en su investigación de un sistema de predicción de precios-venta de inmuebles en el mercado inmobiliario de Chile con el uso de aprendizaje de máquina, tuvo como objetivo general el minimizar el error de predicción del precio en cada tipo de inmueble y en las operaciones inmobiliarias del mercado. El tipo de investigación propuesta es aplicada y tuvo como metodología de desarrollo CRISP-DM enfocándose en el análisis del conocimiento de manera profunda para la solución de problemas. Tuvo como objeto de estudio los precios de los diferentes inmuebles para que sea entrenado la red neuronal, así como las operaciones de inmueble y los resultados arrojados por el modelo. Como resultado de la investigación se comparó el resultado obtenido de la predicción del modelo de red neuronal con el modelo de la empresa general boosted regression y luego calculando la ganancia monetaria obtenida con los mejores modelos desarrollados. Concluyó con que es importante estimar los precios de venta de los inmuebles, teniendo en cuenta la consideración para que la empresa arrendadora pueda generar ingresos en cuestión de la predicción.

(Grajales, 2019) en su investigación de un modelo de predicción de precios de viviendas para el apoyo en la toma de decisiones en las acciones de bienes raíces, tuvo como objetivo general crear un modelo de RNA para la predicción de precios usando machine learning haciendo uso de la obtención de información mediante páginas webs inmobiliarias para la toma de decisiones en la compra informada y venta de propiedades. El tipo de investigación es cuantitativa y uso la metodología CRISP-DM que es muy utilizada en este tipo de proyectos. Tuvo como población y objeto de estudio la información relacionada con los precios de viviendas de las páginas web del municipio de Rionegro. Como resultados tuvo el procesamiento de los datos en un análisis univariable y bivariados, así como la obtención del tipo de modelo a elaborar y los resultados gráficos de la predicción con el valor real del inmueble. Como conclusión, obtuvo que los

modelos que presentaron mejor desempeño fueron gradient boosting machine y random forest, debido a que arrojaron  $R^2$  de 0.75 y 0.77 respectivamente.

(Masaro, y otros, 2017) en su investigación basada en la predicción de precios de lácteos en el sector argentino haciendo uso de redes neuronales artificiales. Tuvo como objetivo ajustar un modelo de redes neuronales para realizar predicciones con precisión de precios de lácteos para apoyar la toma de decisiones en el sector productivo, comparándolos con los resultados de un modelo multivariado. Usó datos del sector desde el 2000 al 2012, construyendo una red con 3 capas, como resultado obtuvo las predicciones comparadas en RMSE: 0.0387 vs. 0.0186; MAD: 0.034 vs. 0.0153, y MPE (%): 7.64 vs. 3.5, correspondiente, apreciando una mejor evolución, destacando también las predicciones de la red con las reales, concluyendo que las RNA entrenadas de una manera eficiente pueden dar mejores resultados que métodos tradicionales de series de tiempo.

En el ámbito nacional se tiene a (Huillca, y otros, 2019) en su investigación con título: Sistema inteligente para la predicción del precio de las acciones mineras en la bolsa de New York basado en un modelo de redes neuronales híbrido y máquina basada en soporte vectorial de regresión, tuvo como objetivo general desarrollar el sistema experto para mejorar la predicción de lo antes mencionado mediante técnicas de machine learning. La metodología se basó en planificación, proceso y el resultado donde obtuvo que mientras mayor es el coeficiente de correlación mejores se muestran los resultados, es por eso que el valor elegido para utilizarse en el modelo, fue el coeficiente que propuso.

En el ámbito local se tiene a (Malaver, 2020) en su investigación Uso de redes neuronales para el pronóstico de ventas en la empresa de Buffets situada en la ciudad de Piura, Perú. Tuvo como objetivo el aplicar redes neuronales para obtener el menor porcentaje de error en la empresa. El tipo de investigación fue aplicada y de diseño cuasiexperimental, donde tuvo como población a las ventas de la empresa misma. Obtuvo como resultado 1% de error con las ventas reales y lo pronosticado por el modelo debido a la buena elección de las variables de entrada para el set de entrenamiento de la red neuronal.

(Zapata, 2019) en su investigación acerca del uso de redes neuronales para determinar pronósticos de inflación en el Perú, tuvo como objetivo determinar los valores por mes pronosticados por la red neuronal, haciendo uso de data histórica entre los años 2008 al 2017 para valores futuros, como resultados incluyó un ajuste de parámetros de la red neuronal a un número máximo de épocas de 2000, un error permitido de convergencia o del resultado de entrenamiento de la red de 0,01; la tasa de aprendizaje de 0,01; y el incremento del aprendizaje en 1,05. Estos datos hicieron que se pronostique los valores futuros de la inflación con un error de 6,9% concluyendo que la base de datos obtenida ayudo a mejorar el error de la predicción, además de la metodología utilizada y los parámetros de la red entrenada.

El desarrollo de la presente investigación se torna en las bases teóricas que ayudarán en el sustento de la misma.

Las redes neuronales artificiales han demostrado ser útiles para modelar las características de distintas situaciones reales. Las ventajas de este enfoque son muchas y la precisión, robustez y rapidez resultantes (Sarhani, y otros, 2021).

Se ha dedicado un gran esfuerzo a desarrollar técnicas para la identificación de sistemas lineales invariantes en el tiempo. La identificación lineal se basa en valores medidos de entrada y salida del sistema. La identificación de los sistemas no lineales se basa también en los valores medidos de entrada y salida, pero, es más difícil (Heidari, y otros, 2019).

La regresión de una variable dependiente,  $Y$ , sobre una variable independiente,  $X$ , es el cálculo del valor más probable de  $Y$  para cada valor de  $X$  a partir de un número finito de mediciones posiblemente ruidosas de  $X$  y los valores asociados de  $Y$ . Las variables  $X$  e  $Y$  suelen ser vectores. En la identificación del sistema, la variable dependiente  $Y$ , es la salida del sistema y la variable independiente,  $X$ , es la entrada del sistema. Para llevar a cabo la identificación del sistema, suele ser necesario asumir alguna forma funcional con parámetros desconocidos. Los valores de los parámetros se eligen para que se ajusten lo mejor posible a los datos observados. En el caso de la regresión lineal, por ejemplo, la salida,  $Y$ , se supone que es una función lineal de la entrada  $X$ , y los parámetros desconocidos,  $a$ , son coeficientes lineales. El enfoque presentado aquí utiliza un método que

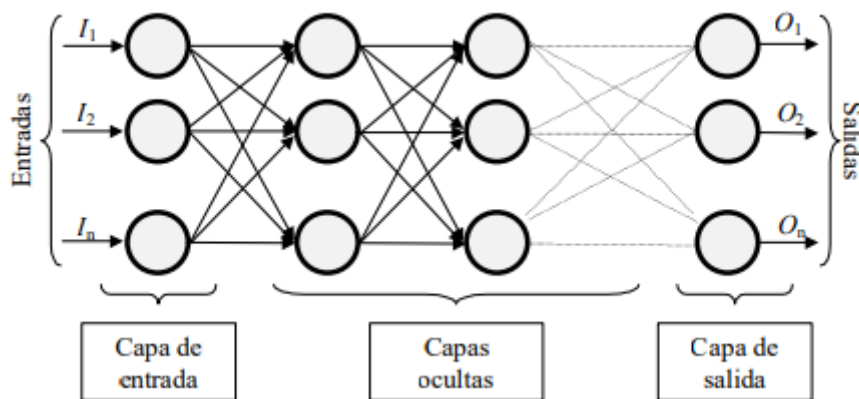
libera de la necesidad de asumir una forma funcional específica (Heidari, y otros, 2019).

Existen varias definiciones para las redes neuronales, desde las más sencillas hasta las más detalladas donde se puede conceptualizar como:

Las RNA son redes interconectadas en gran cantidad y en paralelo con elementos simples estos siendo usualmente adaptativos, y con una organización basada en la jerarquía, las cuales interactúan con el mundo real simulando el funcionamiento del sistema nervioso humano y biológico (Heidari, y otros, 2019).

Elementos básicos que constituyen una red neuronal. En la siguiente Figura 1, se puede ver, un esquema de una red neuronal:

Figura 1: Ejemplo de una red neuronal completamente conectada



La representación está conformada por neuronas interconectadas y construidas en tres capas, las cuales pueden variar. Los datos de entrada ingresan mediante la capa de entrada, pasando a través de la capa oculta y saliendo por la capa de salida. Es importante mencionar que la capa intermedia está compuesta por varias capas (Heidari, y otros, 2019).

Las redes neuronales de avance (FFNN) son la forma más popular forma de modelos de RNA que pueden percibir y aproximar modelos computacionales utilizando su avanzada estructura paralela de capas (Heidari, y otros, 2019).

Se componen de un conjunto de neuronas, que actúan como elementos de procesamiento distribuidos en una serie de capas apiladas totalmente conectadas. El MLP es una clase especial de FFNN. En MLP, las neuronas deben organizarse en un modo unidireccional. La transición de datos en MLP se

produce entre tres clases de capas paralelas: la de entrada, la oculta y las capas de salida (Heidari, y otros, 2019).

#### Análisis predictivo – regresiones

Con la clasificación, el aprendizaje profundo es capaz de establecer correlaciones entre, por ejemplo, los píxeles de una imagen y el nombre de una persona. Esto podría llamarse una predicción estática. De la misma manera, si se expone a suficientes datos correctos, el aprendizaje profundo es capaz de establecer correlaciones entre situaciones del presente y situaciones futuras. Se puede ejecutar una regresión entre estos dos. El evento futuro es como la etiqueta en un sentido. El aprendizaje profundo no necesariamente se interesa por el tiempo, o el hecho de que algo no haya ocurrido todavía (Abu-Naser, y otros, 2019).

#### Preprocesamiento de datos

Para poder construir un modelo de redes neuronales artificiales se necesita un conjunto de datos necesarios para poder predecir el valor de otro. Para el procesamiento de datos se necesita elegir los factores adecuados para el tipo de predicción y eliminar los innecesarios. (Abu-Naser, y otros, 2019).

Además, el conjunto de datos debe contener suficientes muestras para que sean divididas en muestras de entrenamiento y muestras de validación, comúnmente se evalúan los datos y de ser necesario se realiza una normalización para que todos los datos sean reales (Abu-Naser, y otros, 2019).

Para la selección de variables de entrada es necesario en algunos casos depurar los valores que contengan alta correlación, debido a que pueden hacer que el aprendizaje de la red se vea dificultada (Becerra, y otros, 2018).

#### Entrenamiento de una red

El entrenamiento de una red es importante y se basa en utilizar un conjunto de entrenamiento y otro para validar el resultado, estos pueden ser escogidos al azar y pueden ser de distinta calidad, la elección específica puede afectar los resultados globales del entrenamiento y validación. Estos son probados con un algoritmo de entrenamiento, preciso para esperar el comportamiento de la red

neuronal deseado(Machine learning a model for RNA structure prediction, December 2020).

Las neuronas codificadas procesan los parámetros de entrada en dos caminos: haciendo combinaciones lineales para después ejecutar, por medio de una función de activación, esas combinaciones (Smith, 2017).

La esencia de la tasa de aprendizaje proviene de la observación de que el aumento de la tasa de aprendizaje puede tener un efecto negativo a corto plazo y, sin embargo, conseguir un efecto beneficioso a largo plazo. Esta observación conduce a la idea de dejar que la tasa de aprendizaje varíe dentro de un rango de valores en lugar de adoptar un valor fijo o exponencialmente decreciente. Es decir, se fijan unos límites mínimo y máximo y la de aprendizaje varía cíclicamente entre estos límites (Smith, 2017).

#### Técnicas de estimación

Una vez definida la estructura del modelo, es decir, la función paramétrica, y se ha recogido un conjunto de datos, la estimación del parámetro es conceptualmente sencilla: Minimizar la distancia entre las salidas predichas y las salidas medidas (Soltani, y otros, 2019).

#### Visualización de pérdidas y precisión

Después del entrenamiento de una red neuronal, se puede visualizar el progreso de la Red con una función auxiliar personalizada. A medida que una red se entrena, se puede ver que la pérdida disminuye y la precisión aumenta. Como regla general, se busca en un gráfico que las dos líneas converjan o se alineen a medida que aumenta el número de épocas. Si no lo hacen, entonces eso es una señal de que el modelo es inadecuado y se tendrá que volver atrás y cambiar algunos parámetros (Santos , 2020).

#### Fase de validación

La Fase de Prueba o validación, contiene un grupo de datos diferentes a utilizados para el set de entrenamiento, que ayudan a controlar el proceso de aprendizaje de la red. En la fase de entrenamiento se realizarán varios ajustes, donde se presentan situaciones particulares cuando la red se acopla a los valores porque va perdiendo la capacidad de realizar aprendizaje en casos



nuevos, como si se acoplara a los datos que van ingresándose y luego cuando se agregan distintos datos la red no los aprende (Hidalgo, y otros, 2020).

#### Medición del error de predicción

Se van a presentar seguidamente situaciones de errores en el cálculo de una predicción. En la práctica, se busca minimizar ambos tipos de errores optando por el mejor método junto con sets de datos ajustados, y es por eso que existe la medición del error (Díaz, 2020).

#### Uso de métricas

El uso de métricas en el machine learning son proporcionadas en algunos casos, por funciones codificadas de los lenguajes de programación, uno de los más conocidos Python, debido a las librerías disponibles donde se trabaja IA, scikit-learn es una de las librerías donde se utilizan métodos de imputación y evaluación para reproducir resultados de modelos realizados (Yang, y otros, 2022).

#### Score de varianza explicada

La varianza se puede describir como la extensión de los datos si la varianza es cero significa que los datos contienen los mismos valores. La varianza es el promedio de las diferencias cuadradas de la media y se conoce como varianza total de los datos. Después de aplicar un modelo de regresión, la diferencia entre los datos previstos y los reales se denomina varianza residual. La diferencia entre la varianza total y la varianza residual se conoce como varianza explicada. (Shivani, y otros, 2021)

#### Error cuadrático medio estandarizado

El error cuadrático medio estandarizado es la desviación estándar de los valores restantes (error de predicción). Estos valores son una forma de medir que tan lejos esta la línea de regresión de los datos. Por lo que, es una medida que se puede usar para medir cuán dispersos se grafican estos valores. Se puede decir también, que se define en cuán concentrados están los datos en función de la línea de mejor ajuste (Chicco, y otros).

El RMSE se usa ampliamente con situaciones reales tales como el pronóstico, predicción y análisis de regresión. Para validar los resultados resultantes de la experimentación existe la siguiente fórmula:

$$RMSE = \sqrt{(f - o)^2}$$

Donde f representa el valor estimado para los datos que aún no se conocen.

Y o: es el valor obtenido, en otras palabras, datos que se conocen (Wang, y otros, 2018).

Los bienes inmuebles se clasifican comúnmente en urbanos, rústicos y de características o atributos especiales, del cual, la clasificación de este dependerá de la naturaleza del terreno (Rubio, y otros, 2020).

Precio de un inmueble

Un bien está constituido por un grupo de atributos que lo caracterizan, su precio en el mercado inmobiliario se basa en la suma del precio individual de todas estas. Por medio de técnicas de regresión estadística se logra una cuantificación monetaria de todas las propiedades que conforman el inmueble analizado y estiman su valor en el mercado (Picardo, 2019).

Para obtener un precio abierto a mercado más exacto se debe hacer una operación con el valor del terreno, agregado la construcción y obras añadidas.

El valor principal se calcula con el precio del metro cuadrado del terreno, calculándolo con un estudio de mercado, ese valor multiplicado por los metros cuadrados que posee la vivienda. Esos datos se registran en el predio según su clasificación: urbano o rústico (Urbanía, 2016).

De esta manera, es pertinente utilizar un modelo de regresión lineal para predecir el precio de un inmueble tomando como variables a sus características, en su versión más simple, de la siguiente manera:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip} + \epsilon_i$$

donde  $Y_i$  es el precio de la vivienda  $i$ , las  $X_1 \dots X_p$  refieren a los  $p$  atributos de la vivienda  $i$ , los  $\beta_1 \dots \beta_p$  son los coeficientes a estimar y  $E_i$  el término de error, independiente e idénticamente distribuido (iid)  $N(0, \sigma^2)$  (Debray, y otros, 2017).

Estos modelos de predicción de precios de bienes raíces inician en la incógnita de que el precio de un bien se puede dividir en la suma del precio de sus características, o atributos, donde se trata de un tipo de regresión lineal donde sus covariables son todas las características del inmueble y la variable de salida es el resultado, ósea, el precio del inmueble. (Picardo, 2019).

### III. METODOLOGÍA

#### 3.1 Tipo y Diseño de investigación

La presente investigación fue de tipo aplicada porque según el autor (Lozada, 2015) es la investigación que genera conocimiento con su aplicación directa o en un promedio de tiempo en la sociedad, donde presenta un significativo aporte por la utilización del conocimiento proveniente de la investigación básica.

Asimismo, el enfoque de la presente investigación fue cuantitativa, debido a que se analizaron resultados por medio de la estadística y como indica los autores (Fernández, y otros, 2002) es la investigación en la cual se recopilan y se analizan datos sobre variables, trata de dar sentido a los objetivos de la misma con los resultados a través de las muestras para hacer la consecuencia a una población de la cual toda la muestra procede.

La investigación presente fue de nivel explicativo, debido a que se explicó cómo va desarrollándose y simulando un modelo de red neuronal para la predicción de un precio real, también (List, 2018) la investigación explicativa orienta a los investigadores a estudiar y entender el entorno de manera más profunda, para obtener resultado de manera eficiente.

Por último, el diseño de la investigación presentada fue Preexperimental, debido a que se sometió el modelo a varias pruebas con los valores obtenidos de los inmuebles para poder medir el grado de efectividad (Diseños Preexperimentales y cuasiexperimentales aplicados a las ciencias sociales y educación, 2020) es un diseño donde se utilizan preexperimental para acercarse más al fenómeno que se está investigando, utilizando un estímulo para generar una hipótesis y posterior medir sus variables para observar sus efectos.

#### 3.2 Variables y Operacionalización

La presente investigación contó con dos variables, la variable independiente es el modelo basado en redes neuronales con definición: Un sistema computarizado integrado por un gran conjunto de elementos simples, muy interconectados, de los cuales se procesa información debido a su estado dinámico en función de entradas y salidas (Heidari, y otros, 2019), y la variable dependiente Predicción

del precio de un inmueble que se define como la predicción es el anuncio de lo que se espera que puede suceder. Es un elemento importante de las ciencias, en general, pues permiten iniciar experimentos y contrastar el hecho esperado con la realidad. (Westreicher, 2020). Los bienes inmuebles son aquellos que como su nombre lo indica son inmóviles, que no se pueden separar del terreno en el que se encuentran, pero de manera más específica de los bienes raíces, de tal modo, son terrenos o construcciones que se caracterizan como bienes físicos o jurídicamente. (Téllez, y otros, 2021). La variable independiente tiene cuatro dimensiones las cuales son: Entrenamiento de la red, parámetros y patrones, modelo y criterios de evaluación. Para la primera dimensión tenemos el indicador Incremento de la precisión, para la dimensión modelo tenemos el indicador Algoritmo, para el indicador criterios de evaluación tenemos porcentaje mínimo de la predicción. Para la variable dependiente tenemos las dimensiones: Inmuebles departamentos en la ciudad de Piura, Inmuebles tipo casa en la ciudad de Piura, para la primera dimensión está el indicador precio predicho y para la segunda dimensión está el indicador precio predicho igualmente. La Operalización se detalla más en Anexo.

### 3.3 Población, muestra y muestreo

(Artigas, y otros, 2010) indican que la población es un grupo finito o infinito de componentes con características similares que servirán para dar una conclusión a una investigación. En la siguiente investigación para medir la efectividad del modelo de red neuronal se enfoca como unidad de análisis el mismo modelo, siendo una población finita los registros de la predicción del precio del inmueble, del error de predicción y del incremento de la precisión estableciendo la muestra elegida por conveniencia por los investigadores, cantidad pertinente para analizar las variables.

### 3.4 Técnicas e instrumentos de recolección de datos

La técnica utilizada para la investigación fue la observación, utilizada para registrar las distintas simulaciones de la red neuronal modelada, y como indica (Simister, 2017) en su forma más simple, la observación implica ver cosas, como objetos, procesos, relaciones, acontecimientos, y registrar formalmente la información.

Consistió en aplicar el instrumento ficha de análisis documental para identificar el modelo más adecuado para desarrollar una predicción de precios de inmuebles.

Seguidamente se aplicó el instrumento de ficha de registro que según (Hernández, y otros, 2020) es un instrumento que contiene los aspectos diseñados a recolectar, esta debe ser sencilla y de fácil entendimiento para que el investigador pueda completarla. Utilizada para evaluar el incremento de la precisión de una red neuronal, basado en las características que harán que varíe el resultado obtenido de la predicción tales como los datos procesados, el set de datos, el número de neuronas y capas y el ciclo de entrenamiento.

Se aplicó otra ficha de registro para evaluar el porcentaje mínimo de error de la predicción después del ajuste de la red neuronal, para verificar cual obtiene un acercamiento más al precio real.

Para evaluar el precio predicho se utilizó el instrumento de ficha de registro para obtener el resultado del precio predicho por la red neuronal y el precio real que se quiso predecir.

Para la presente investigación se tomó consideración avalar la validez de las fichas de registro, debido a que el instrumento se elaboró únicamente para recolectar información necesaria con los ítems elaborados para la investigación, todo esto con la finalidad de evaluar los objetivos inicialmente planteados y (Almanasreh, y otros, 2019) argumentan que la validez del contenido proporciona pruebas sobre el grado en que los elementos de un instrumento de evaluación son relevantes y representativos del contenido de un instrumento de evaluación es relevante y representativo del constructo que se pretende evaluar.

La confiabilidad de la investigación como indica (Roberts, y otros, 2006) son formas de demostrar y comunicar el rigor de los procesos de investigación y la fiabilidad de los resultados de la investigación. Esta depende de una serie de características de la investigación: la pregunta inicial de la investigación, cómo se recogen los datos, incluyendo cuándo y de quién, cómo se analizan y qué conclusiones se extraen, concluyendo el uso del test retest.

Tabla N° 1: Técnicas e instrumentos

<b>Indicadores</b>	<b>Técnica</b>	<b>Instrumentos</b>	<b>Fuente</b>	<b>Informante</b>
Incremento de la precisión	Observación	Ficha de registro	creadas por la aplicación	Investigador
Porcentaje de error mínimo de la predicción	Observación	Ficha de registro	creadas por la aplicación	Investigador
Precio predicho	Observación	Ficha de registro	creadas por la aplicación	Investigador
Parámetros de entrada	Observación	Ficha de registro	creadas por la aplicación	Investigador
Modelo	Análisis documental	Ficha de análisis documental	Documentos	Investigador

Fuente: Elaboración propia

### 3.5 Procedimientos

Las fichas de registro buscan medir la efectividad del modelo de red neuronal desarrollado para la predicción de un precio. Se aplicó la revisión para elegir el modelo preciso para la elaboración de la red neuronal, luego se registró en la debida ficha los datos de la simulación de cada resultado dado por el modelo, donde se podrá medir el número de neuronas de entrada (los datos de entrada), las capas ocultas que contienen los epochs o épocas de cómo se va a entrenar la red, el set de entrenamiento que contiene un número de datos específicos obtenidos mediante otra ficha de registro para los parámetros, es decir, los datos que servirán para que la red neuronal aprenda de estos y pueda predecir, esto para evaluar la eficiencia en la predicción. Por último, se registró el resultado de la predicción del precio de un inmueble ingresado y el real en la red neuronal para evaluar el error entre los resultados.

### 3.6 Métodos de análisis de datos

El procesamiento de datos se hizo mediante la obtención de características de inmuebles de páginas inmobiliarias de la ciudad de Piura, los resultados arrojados por el modelo desarrollado se analizaron con estadística descriptiva y de regresión lineal para comparar los precios reales con los precios predichos y analizar la función de pérdida del entrenamiento de la red mediante gráficos lineales, que permita describir los resultados de cada configuración y para evaluar el error normal, error absoluto, el error cuadrático medio y el score de

varianza que fue hallado mediante una función codificada de la red neuronal, para medir la efectividad del modelo.

### 3.7 Aspectos éticos

Para desarrollar la actual investigación se tomaron en cuenta los aspectos éticos tales como: respetar toda base teórica, recurriendo a los antecedentes previamente planteados, citando correctamente a autores y ciñéndose a la ejecución de los instrumentos para ser evaluados. Resguardar toda integridad de los datos que han intervenido en la investigación, siendo el caso de ser nombrados siempre y cuando exista una autorización.



#### IV. RESULTADOS

Para el objetivo **Determinar el modelo a desarrollar para la red neuronal**, se obtuvo según (Hilera, y otros, 2000) definiendo un modelo secuencial de única entrada y salida basado en ajustes ReLu (rectified linear unit) un modelo de red neuronal basado en 4 fases.

Las fases de este modelo comprenden su metodología, tanto para su preparamiento, desarrollo y validación, las cuales son:

- Definición de la red neuronal: Donde se define toda la configuración de la red neuronal, set de datos, modelo, neuronas, capas, función de activación.
- Entrenamiento: Donde se define que tipo de aprendizaje se va a utilizar, para el modelo secuencial fue el supervisado.
- Validación: Donde se pone a prueba el modelo entrenado y se hace la ejecución de la red neuronal,
- Mantenimiento: Donde luego se realiza un mantenimiento que significa, el volver a validar el modelo y ajustar sus configuraciones para que se mantengan en el tiempo.

Para el objetivo **Identificar los parámetros de entrada para el modelo de la predicción de los precios**, se obtuvo la recopilación de datos mediante web scrapping de distintas webs que poseían datos de inmuebles en venta en la ciudad de Piura obteniendo como valores máximos mostrados en la siguiente tabla:

Tabla N° 2: Valores obtenidos parámetros de entrada en Piura

	<b>Año</b>	<b>Tipo de Inmueble: Casas</b>	<b>Páginas utilizadas</b>
<b>Valores Máximos</b>	<b>1</b>	<b>80</b>	<b>3</b>

Fuente: Elaboración propia

Para complementar los datos obtenidos dentro de la ciudad de Piura también se utilizó esta herramienta para conseguir datos sobre inmuebles valorizados por métodos estadísticos y ya vendidos, obteniendo como valores los mostrados en la siguiente tabla:

Tabla N° 3: Valores obtenidos parámetros de entrada pág. alternas

	Año	Tipo de Inmueble: Casas	Páginas utilizadas
Valores Máximos	3	21000	1

Fuente: Elaboración propia

Posteriormente, se hizo el procesamiento de estos datos y normalización para formar parte del set de entrenamiento de la red neuronal, los cuales serían los parámetros de entrada de la misma. Para esto se analizaron los datos mediante un gráfico de dispersión para visualizar la relación entre dos variables, siendo estas el precio del inmueble y los m2 habitables del terreno:

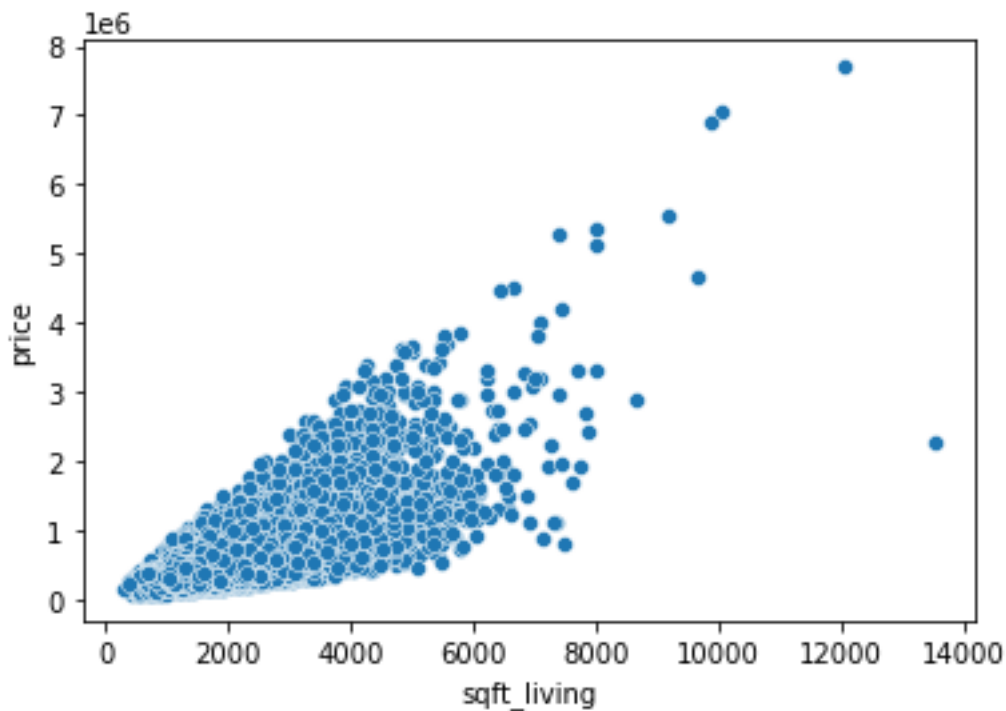


Figura 3: Diagrama de Dispersión de Inmuebles

Fuente: Elaboración Propia

Para el objetivo **Determinar la efectividad del modelo de redes neuronales en la predicción de un inmueble**, realizó la configuración para el entrenamiento de la red neuronal y analizar el resultado del mismo. Como primera configuración los datos se pueden apreciar en la siguiente tabla:

Ficha de registro: Resumen de red neuronal construida y entrenada N° 1	
Tipo de prueba: test	Indicador: Aumento de la precisión.
Datos procesados	habitaciones', 'baños', 'm2_habitables', 'm2_terreno', 'pisos', 'm2_techado', 'm2_sotano' - 7
Tipo de variables	numéricas
Tamaño del set de entrenamiento	0.25
Número de Epochs	100
Número de neuronas	200
Número de Capas	4
Función de pérdida	mean_squared_error

Tabla N° 4: Resumen de red neuronal construida y entrenada N°1

Fuente: Elaboración propia

El historial del entrenamiento y la pérdida del entrenamiento durante las epochs esta resumida en la siguiente tabla:

Tabla N° 4: Resumen de historial de entrenamiento

	loss	val_loss
0	0.001213	0.001425
1	0.001213	0.001425
2	0.001096	0.001468
3	0.001071	0.001226
...	...	...
97	0.000890	0.000995
98	0.000896	0.000998
99	0.000908	0.001020

Fuente: Elaboración propia

Estos datos fueron representados mediante la función `history.keys()` y posteriormente representados en el siguiente gráfico:

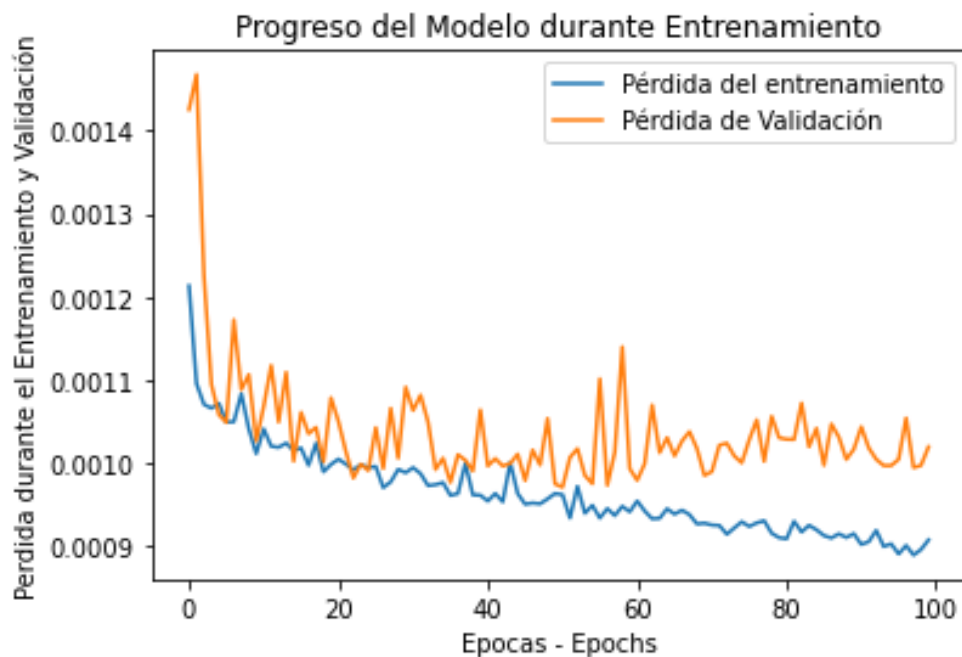


Figura 4: Gráfico rendimiento del modelo durante el entrenamiento N°1

Fuente: Elaboración Propia

El gráfico de curvas de aprendizaje muestra el progreso del modelo durante el entrenamiento, representado por los valores de pérdida del entrenamiento y validación. Un buen ajuste se identifica si la curva de pérdida de entrenamiento disminuye hasta conseguir un punto de estabilidad y si consigue tener una brecha pequeña con la curva de validación, en este caso se usó un `test size = 0.25` y un `validation Split = 0.5` por lo cual se muestran movimientos ruidosos que pueden demostrar que el set de datos de validación es escaso o reducido.

Se puso a prueba el modelo con la predicción del precio de una casa de ocasión en venta donde se puede apreciar:

Tabla N° 5: Resultado Simulación N°1

Simulación	Resultado Predicción	Valor de validación
1	231103	140000

Fuente: Elaboración propia

Para interpretar el error obtenido del entrenamiento se realizó la siguiente tabla:

Tabla N° 6: medición del error del modelo N°1

Mae (y_test, prediction)	0.020651905247304213
Mse (y_test, prediction)	0.0010650129809143634
Explained variance score (y_test, prediction)	0.5532307782516434

Fuente: Elaboración propia

En esta tabla podemos apreciar un variance score de 0.55 lo cual indica que no es tan adecuado el modelo entrenado ya que ha conseguido una puntuación lejana de 1,0 la cual es adecuada para el modelo.

Para la segunda configuración, los datos del modelo se detallan en la siguiente tabla:

Tabla N° 7: Resumen de red neuronal construida y entrenada N°2

<b>Ficha de registro: Resumen de red neuronal construida y entrenada N° 2</b>	
<b>Tipo de prueba: test</b>	<b>Indicador: Aumento de la precisión.</b>
Datos procesados	habitaciones', 'baños', 'm2_habitables', 'm2_terreno', 'pisos', 'm2_techado', 'm2_sotano' - 7
Tipo de variables	numéricas
Tamaño del set de entrenamiento	0.3
Número de Epochs	500
Número de neuronas	53
Número de Capas	4
Función de pérdida	mean_squared_error

Fuente: Elaboración propia

Se decidió por aumentar el número de epochs para que el entrenamiento resulte más estable y se disminuyó el número de neuronas también aumentando el set de entrenamiento y validación, los resultados del modelo se graficaron en la siguiente figura:

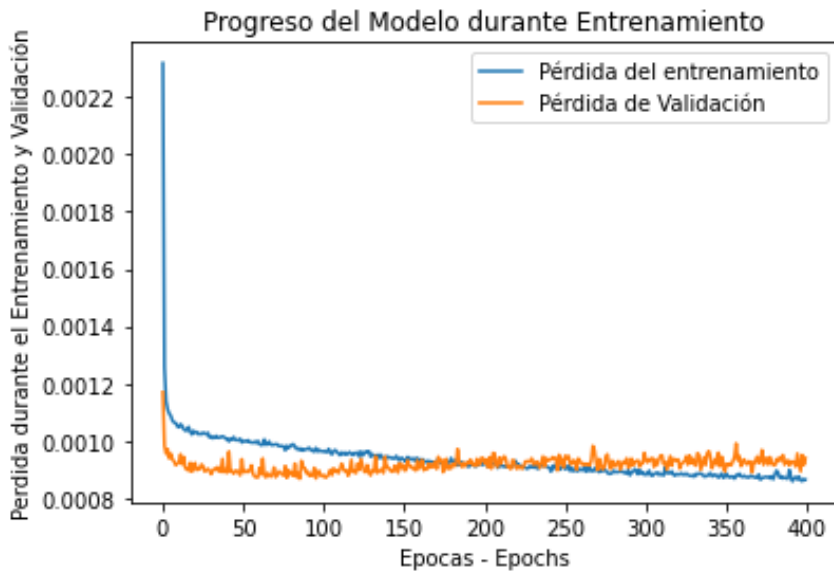


Figura 5: Gráfico rendimiento del modelo durante el entrenamiento N°2

Fuente: Elaboración Propia

En el gráfico se puede mostrar que la pérdida se estabilizó considerablemente sin embargo la brecha entre las dos curvas no se aprecia muy bien, se utilizó un `test_size = 0,3` y un `validation_split = 0,4` y a pesar del aumento de epochs no se obtuvo un buen resultado del entrenamiento, dando la tabla de medición del error de la siguiente manera:

Tabla N° 8: Resultado de Simulación N°2

Simulación	Resultado Predicción	Valor de validación
2	313780	185000

Fuente: Elaboración propia

Tabla N° 9: medición del error del modelo N°2

Mae (y_test, prediction)	0.020369810766150073
Mse (y_test, prediction)	0.0011016155079383104
Explained variance score (y_test, prediction)	0.5455105675913179

Fuente: Elaboración propia

A pesar de que el entrenamiento ha sido más estable, el variance score ha disminuido a 0.54 lo cual no lo hace aceptable para predecir precio de inmuebles

Para la tercera configuración, los datos del modelo se detallan en la siguiente tabla:

Tabla N° 10: Resumen de red neuronal construida y entrenada N°3

<b>Ficha de registro: Resumen de red neuronal construida y entrenada N° 3</b>	
<b>Tipo de prueba: test</b>	<b>Indicador: Aumento de la precisión.</b>
Datos procesados	habitaciones', 'baños', 'm2_habitables', 'm2_terreno', 'pisos', - 5
Tipo de variables	Numéricas
Tamaño del set de entrenamiento	0.3
Número de Epochs	400
Número de neuronas	53
Número de Capas	4
Función de pérdida	mean_squared_error

Fuente: Elaboración propia

Se decidió por disminuir el número de epochs sin modificar el número de neuronas y capas, los resultados del modelo se graficaron en la siguiente figura:

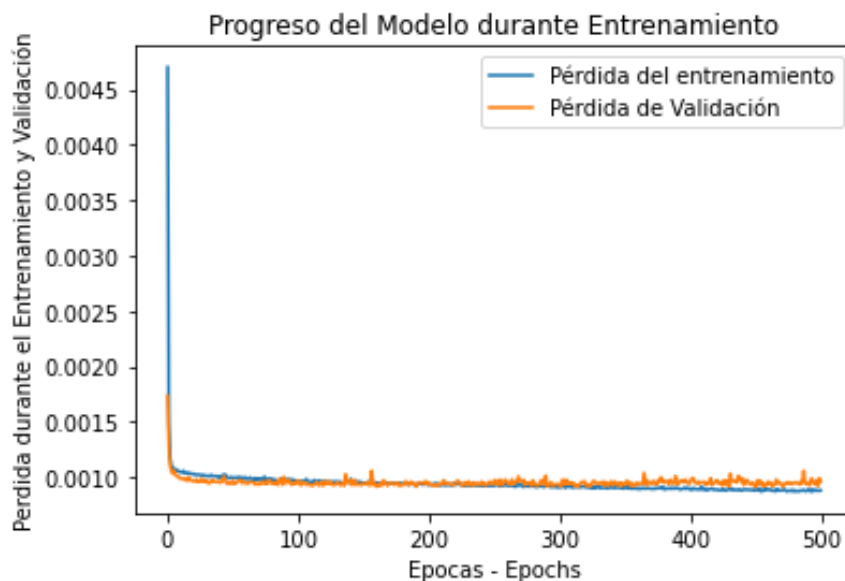


Figura 6: Gráfico rendimiento del modelo durante el entrenamiento N°3

Fuente: Elaboración Propia

El grafico muestra una estabilidad entre la perdida de entrenamiento y la perdida de validación debido a que se empleó una menor cantidad de epochs manteniendo el test size y el validation Split, agregando un batch\_size = 128 y un random state = 101 los cuales ayudaron a estabilizar el entrenamiento del modelo. La tabla de la medición de errores es la siguiente:

Tabla N° 11: Resultado de Simulación N°3

Simulación	Resultado Predicción	Valor de validación
3	196146	140000

Fuente: Elaboración propia

Tabla N° 12: medición del error del modelo N°3

Mae (y_test, prediction)	0.010912378749528556
Mse (y_test, prediction)	0.000920310691491143
Explained variance score (y_test, prediction)	0.7998036653422818

Fuente: Elaboración propia

En esta tabla se puede apreciar un mejor variance score de 0.8, el deseado es 1,0 pero 0.8 es un score aceptable obtenido para la predicción de inmuebles en base a sus características.

Para el objetivo específico: **Evaluar el porcentaje de error mínimo para que la red neuronal sea efectiva**

Para la evaluación del error en la predicción del resultado del modelo, se tiene la siguiente tabla:

Tabla N° 13: Resultados de predicción del modelo

SIMULACIÓN	PREDICCIÓN	VALOR REAL	ERROR NORMAL	ERROR ABSOLUTO
1	231103	140000	-65%	65%
2	243780	185000	-31%	31%
3	176146	140000	-25%	25%

Fuente: Elaboración Propia

En esta tabla podemos apreciar como el porcentaje de error disminuyó con los ajustes en las simulaciones del modelo de red neuronal, obteniendo como valor mínimo un error del 25% en la tercera simulación con un score de varianza de 0.80.



## V. DISCUSIÓN

El resultado obtenido al procesar los datos recolectados relacionados al objetivo, Determinar el modelo a desarrollar para la red neuronal, se puede afirmar que el modelo de red neuronal adecuado es el secuencial de única entrada y salida, debido a que la red neuronal construida necesitaba de los valores de entrada para poder predecir un único valor, el precio de un inmueble. Estos hallazgos difieren de la investigación de (Masaro, y otros, 2017) que explica que, para la predicción de valores que se extienden en el tiempo se necesita de un modelo perceptrón multicapa, la composición de la red neuronal es similar, pero la salida de la red neuronal llega a ser de múltiples valores, debido a que se aplica en series de tiempo. Sin embargo, la manera en la que se ejecuta el entrenamiento de la red neuronal en los dos casos es igual, un aprendizaje supervisado, donde ya se tiene conocimiento del set de entrenamiento y por lo tanto se aplica de manera directa, a diferencia del no supervisado donde no se saben los patrones y se van hallando con el tiempo.

Distintos autores como (Grajales, 2019) y (Guerrero, 2018) de las investigaciones estudiadas, concuerdan en que para desarrollar una red neuronal para la predicción de un único valor se necesita de un modelo secuencial, porque la RNA necesita reemplazar el método estadístico tradicional basado en regresión lineal, utilizado para la predicción de precios de inmuebles.

De acuerdo al objetivo Identificar los parámetros de entrada para el modelo de la predicción de los precios, los resultados obtenidos mediante el análisis de las características de los inmuebles, se tomaron 7 valores, incluyendo como esenciales los metros cuadrados del inmueble y del terreno, esto concuerda con la investigación realizada por (Bozanic, 2020) donde el análisis de datos correspondiente dio como resultado la selección de 7 valores, siendo de estos igualmente los más importantes los ya mencionados. El estudio también relaciona las habitaciones, los pisos y los baños como parte fundamental de las características de un inmueble en su estimación. En los resultados hallados de la presente investigación se usan estos valores para la predicción de un inmueble en venta, sin embargo, (Bozanic, 2020) indica que se utilizan estos datos para

dos tipos de operación, usando datos de casas, departamentos y oficinas en la venta y el arriendo.

Para la selección de las variables que incluyeron el set de entrenamiento de la red neuronal se necesitó cambiar todos los datos recolectados del instrumento a tipo numérico para que la red neuronal haya sido más efectiva. Se tomaron rangos de 1-5 para indicar la condición y el uso del código postal para referirse a la zona de ubicación del inmueble, este análisis se puede contrastar con la investigación de (Grajales, 2019) que indica en su investigación que esos datos fueron descartados después de la limpieza de datos, debido a que iba a restar precisión en la predicción de la red neuronal de regresión lineal, sin embargo, en su investigación señala que no solo se usa este modelo de RNA sino 5, donde en los demás si se toman en cuenta estos valores, esto a diferencia de la presente investigación en la cual se usa el modelo secuencial y se toman estos datos en un inicio obtenidas del instrumento de recolección de datos.

Con respecto al objetivo planteado Determinar la efectividad del modelo de redes neuronales en la predicción de un inmueble, en base a sus características, existen distintas maneras de configurar una red neuronal para la predicción de precios, la efectividad de una red neuronal se basa tanto en los resultados de su entrenamiento como en medir el error del valor aproximado y real, analizando las investigaciones, (Malaver, 2020) indica que hizo uso de 3 configuraciones basadas en un set de datos, epochs y numero de capas y neuronas, esto concuerda con los resultados obtenidos, debido a que se necesita simular cada configuración de la arquitectura de la red para obtener el error mínimo. Esta técnica propone unos resultados basados en graficar el proceso de pérdida durante el modelo, tal y como lo señala (Santos , 2020), esto desenlaza en métricas para evaluar el desempeño del mismo. Se emplearon las métricas de score de varianza y RMSE, este último, también es empleado en la investigación de (Huillca, y otros, 2019) utilizado para la predicción de precio diario de la bolsa de valores, esta métrica es importante para evaluar situaciones donde se empleen regresiones y es por eso que se usó en la investigación.

Como último punto, del objetivo Evaluar el porcentaje de error mínimo para que la red neuronal sea efectiva se pudo rescatar que el error en la predicción del

precio de un inmueble influyen distintos factores como las variables, la configuración de la red y su posterior evaluación para seguir realizando los ajustes simulación tras simulación hasta obtener el error mínimo, a esto se tuvo que la red neuronal arrojó resultados que dieron como error mínimo un 25%, después de dos simulaciones donde se obtuvo un 65% en la primera configuración y un 35% en la segunda configuración, uno de los parámetros utilizados para darle estabilidad al modelo y pueda ser efectivo fue el random state, un parámetro del aprendizaje supervisado, y el test size (tamaño del conjunto de prueba), así como el seleccionar ciertas variables que son más determinantes para entrenar el modelo y descartar otras, este análisis se puede comparar con la investigación de (Malaver, 2020) donde elaboró simulaciones en la predicción de un pronóstico de ventas donde obtuvo 14%, 11% y 1% de error, considerando este último como el mínimo, la manera en la que se obtuvo este resultado para ser considerada efectiva fue en el uso de la métrica de señal de rastreo, usada para pronósticos y la cual le permitió evaluar si el modelo era el adecuado, a comparación de la presente investigación donde se usó el score de varianza importada desde sklearn el cual indicaba una aproximación a 1,0 lo cual se interpreta como mejor puntuación.

Obtener un resultado de error mucho menor se considera más complicado con respecto a la investigación de (Malaver, 2020) debido a que esta solo utiliza 5 variables que se centran en la fecha de venta, tipo de servicio, y el total de la venta, que pueden ser fácilmente utilizados en su totalidad como set de entrenamiento y que su variable determinante es el total de la venta, además de usar un modelo que es el más utilizado para los pronósticos, ayudó en la elaboración del modelo, en el caso de las regresiones existen múltiples modelos que pueden ser utilizados para obtener los datos relevantes y decisivos sobre la predicción y hay una amplia variedad de análisis, de la cual se escogió solo una, el modelo secuencial de RNA.

## VI. CONCLUSIONES

En la presente investigación, se llegó a la conclusión de que el uso de un modelo secuencial es el adecuado para desarrollar un modelo de redes neuronales basado en la predicción, donde se demostró un funcionamiento adecuado en función de entradas de datos, única salida y en su metodología de construcción, que la hace adecuada para los modelos basados en regresión.

Queda demostrado que los parámetros de entrada asignados son los correctos para desarrollar la red neuronal, esto debido a que características como los metros cuadrados del inmueble, las habitaciones, baños y número de pisos son variables cruciales en la estimación de un bien, a pesar de haber descartado algunos datos como la vista, condición y el código postal en el procesamiento de datos, mostraron coherencia al ser utilizados tanto como set de entrenamiento en las simulaciones y como parámetros para la predicción.

Para la determinación de la efectividad de un modelo de redes neuronales en la predicción de inmuebles se concluyó que el modelo es efectivo no solo por las correctas configuraciones realizadas tras cada simulación, sino por el uso de métricas como el score de varianza, que permitieron evaluar el modelo en base a su precisión durante el entrenamiento, siendo efectivo con un score de 0.8, aceptable para los parámetros utilizados y el número de épocas empleadas en el último registro del proceso de modelamiento.

Por último, se demostró que el porcentaje mínimo para que la red neuronal sea efectiva es de 25%, gracias a la visualización de pérdida del modelo, donde se pudo analizar y ajustar la configuración en cada registro, además de que fue mejorando de manera sucesiva a pesar de no haber procesado el parámetro de ubicación de un inmueble, un valor más complicado de evaluar en valores numéricos.

## VII. RECOMENDACIONES

- Se recomienda a los investigadores la elaboración de investigaciones con más visión a los indicadores que puedan evaluar un modelo de redes neuronales, ya que con los avances tecnológicos serán más complejos, a fin de lograr unos resultados mucho más precisos.
- Asimismo, tomar en cuenta más parámetros para la creación de un set de entrenamiento, para mejorar el porcentaje de error en la predicción en la red neuronal. Además, se recomienda crear un método que permita ingresar datos periódicamente en las redes neuronales de predicción, esto para evitar que el modelo pierda efectividad con el tiempo.
- Se recomienda a las empresas y agencias inmobiliarias, usar inteligencia artificial para poder predecir futuras propiedades y así tomar decisiones que puedan sobreponerse frente a la competencia en el sector inmobiliario.

## REFERENCIAS

**Abu-Naser, Samy y M. Barhoom, Alaa. 2019.** *Energy Efficiency Prediction using Artificial Neural Network*. s.l. : International Journal of Academic Pedagogical Research, 2019.

**Almanasreh, Enas, Moles, Rebekah y Chen, Timoty F. 2019.** Evaluation of methods used for estimating content validity. [aut. libro] Enas Almanasreh, Rebekah Moles y Timoty F. Chen. *Research in Social and Administrative Pharmacy*. Australia : The University of Sydney, 2019, Vols. 15, Issue 2, págs. 214-221.

**Artigas, Wileidys y Robles, Miguel. 2010.** *Metodología de la Investigación: Una discusión necesaria en Universidades Zulianas*. Mexico : Revista Digital Universitaria, Dirección General de Cómputo y de Tecnologías de Información y Comunicación - UNAM, 2010. ISSN: 1067-6079.

**Becerra, Miguel, y otros. 2018.** *Redes Neuronales en predicción de mercados financieros: una aplicación en la bolsa mexicana de valores*. México : Tecnológico Nacional de México en Celaya, 2018. Pistas Educativas Vol. 40 - ISSN: 2448-847X.

**Bozanic, Mirko Slovan. 2020.** *SISTEMA DE PREDICCIÓN DE PRECIOS-VENTA DE INMUEBLES EN EL MERCADO DEL SECTOR INMOBILIARIO DE LA REGIÓN METROPOLITANA DE LA REPÚBLICA DE CHILE CON EL USO DE ALGORITMOS DE MACHINE LEARNING*. Región Metropolitana de Santiago : Universidad de Chile - Facultad de Ciencias Físicas y Matemáticas - Departamento de Ingeniería Industrial, 2020. Disponible en <http://repositorio.uchile.cl/handle/2250/176909>.

**Chicco, Davide, Warrens, Matthijs y Jurman, Giuseppe.** *The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation*.

**Debray, Thomas, y otros. 2017.** *A guide to systematic review and meta-analysis of prediction model performance*. s.l. : bmj, 2017. vol. 356.

**Díaz, Jorge. 2020.** *Precisión del pronóstico de la propagación del covid-19 en Colombia*. Colombia : Secretaría de Educación de Soacha, Cundinamarca., 2020. ISSN: 0121-7372.

*Diseños Preexperimentales y cuasiexperimentales aplicados a las ciencias sociales y educación.* **Chávez, Sarah, Esparza, Oscar y Riosvelasco, Leticia. 2020.** 2, Ciudad de México, México : Enseñanza e Investigación en Psicología, 2020, Vol. 2. 167-178.

**Fernández, Pita y Díaz, Pértegas. 2002.** *Investigación cuantitativa y cualitativa*. Coruña, España : Unidad de Epidemiología Clínica y Bioestadística. Complejo Hospitalario-Universitario Juan Canalejo., 2002.

**García, Luis Alberto y Matienzo, Hector Wildoro. 2020.** *Reactivación del sector inmobiliario considerando el factor social y económico*. Lima, Perú : Escuela Profesional de Ingeniería Civil, Facultad de Ingeniería, Universidad Ricardo Palma, 2020. URL: <http://repositorio.urp.edu.pe/handle/URP/3829>.

**Grajales, Yuri. 2019.** *Modelo de predicción de precios de viviendas en el municipio de Rionegro para apoyar a la toma de decisiones de Compra y Venta de Propiedad Raíz.* Medellín, Colombia : Universidad Pontificia Bolivariana, 2019.

**Guerrero, Carlos Alberto. 2018.** *Modelo para determinar la existencia del fenómeno denominado Burbuja Inmobiliaria en el Perú.* Lima, Perú : Anales Científicos, National Agrarian University, 2018. ISSN-e 2519-7398.

**Heidari, Ali, y otros. 2019.** *An efficient hybrid multilayer perceptron neural network with grasshopper optimization.* Australia : Soft Comput, 2019. DOI: <https://doi.org/10.1007/s00500-018-3424-2>.

**Hernández, Sandra y Duana, Danae. 2020.** Técnicas e instrumentos de recolección de datos. *Boletín Científico de las Ciencias Económico Administrativas del ICEA.* [En línea] 5 de diciembre de 2020. [Citado el: 19 de noviembre de 2021.] <https://repository.uaeh.edu.mx/revistas/index.php/icea/article/view/6019/7678>. ISSN: 2007-4913.

**Hidalgo, Iván, y otros. 2020.** Estudio comparativo de los algoritmos de backpropagation (BP) y multiple linear regression (MLR) a través del análisis estadístico de datos aplicado a redes neuronales artificiales. *Revista boletín REDIPE.* [En línea] 1 de marzo de 2020. [Citado el: 19 de noviembre de 2021.] <https://revista.redipe.org/index.php/1/article/view/939>. Vol. 9 Núm. 3.

**Hilera, José y Martínez, Víctor. 2000.** *REDES NEURONALES ARTIFICIALES.* Madrid, España : Alfaomega, 2000.

**Huillca, Jessica y Quispe, Renzo. 2019.** *Sistema inteligente para la predicción del precio diario de las acciones mineras en la Bolsa de Valores de New York usando un modelo híbrido de redes neuronales y máquina de soporte vectorial de regresión.* Lima, Perú : Universidad Nacional Mayor de San Marcos, Facultad de Ingeniería de Sistemas e Informática, 2019. Consultado: 07 de octubre de 2021. Disponible en: <https://cybertesis.unmsm.edu.pe/handle/20.500.12672/11318>.

**List, Christian. 2018.** *Levels: Descriptive, Explanatory, and Ontological. Noûs.* New Jersey, United States : Wiley Periodicals, Inc., 2018.

**Lozada, Jose. 2015.** *Investigación Aplicada: Definición, Propiedad Intelectual e Industria.* Ambato, Quito : Revista de divulgación científica de la Universidad Tecnológica Indoamérica, 2015. ISSN-e 1390-9592.

*Machine learning a model for RNA structure prediction.* **Calonaci, Nicola, y otros. December 2020.** 4, s.l.: NAR Genomics and Bioinformatics, December 2020, Vol. 2. Disponible en: <https://doi.org/10.1093/nargab/lqaa090>.

**Malaver, Manuel Alexander. 2020.** *APLICACIÓN DE REDES NEURONALES PARA DETERMINAR EL PRONÓSTICO DE LAS VENTAS EN LA EMPRESA CATERING & BUFFETS MyS UBICADA EN LA CIUDAD DE PIURA.* Piura : Universidad César Vallejo, 2020.

**Masaro, Jimena y García, Rodrigo. 2017.** *Redes Neuronales para la predicción de precios en el sector lácteo argentino*. Buenos Aires, Argentina : Facultad de Ciencias Económicas, Cuadernos del CIMBAGE, 2017. ISSN: 1666-5112.

**Picardo, Pablo. 2019.** *Predicción de precios de vivienda, Aprendizaje estadístico con datos de oferta y transacciones para la ciudad de Montevideo*. Montevideo, Uruguay : Documento de trabajo del Banco Central del Uruguay 002-2019, 2019. ISSN 1688-7565.

**Piura News. 2015.** PIURA ES UNA CIUDAD ALTAMENTE ATRACTIVA PARA EL SECTOR INMOBILIARIO. <http://futurasolucionesinmobiliarias.pe/>. [En línea] Revista Piura News, Mayo de 2015. [Citado el: 21 de octubre de 2021.] <http://futurasolucionesinmobiliarias.pe/piura-es-una-ciudad-altamente-atractiva-para-el-sector-inmobiliario/>.

**Roberts, Paula y Priest, Helena. 2006.** *Reliability and validity in research*. London, United Kingdom : Royal College of Nursing Publishing Company (RCN), 2006. Vol. 20, Issue 44.

**Rouhiainen, Lasse. 2018.** *Inteligencia artificial, 101 cosas que debes saber hoy sobre nuestro futuro*. Barcelona, España : Alienta Editorial, 2018. ISBN: 978-84-17568-08-5.

**Rubio, Clara y Martínez, Joaquín. 2020.** *La base imponible del Impuesto sobre Bienes*. Zaragoza, España : Facultad de Derecho de Zaragoza, 2020.

**Santos , Marco. 2020.** towardsdatascience. *I Tried Using Deep Learning to Predict the Stock Market - Forecasting Stock Prices with Neural Networks*. [En línea] 3 de mayo de 2020. [Citado el: 16 de octubre de 2021.] <https://towardsdatascience.com/using-deep-learning-ai-to-predict-the-stock-market-9399cf15a312>.

**Sarhani, Rekia, Saddek, Hafsa y Ghaitaoui, Touhami. 2021.** Modélisation et simulation d'une cellule Photovoltaïque organique par réseau neuronale artificiel. *dSPACE - UNIVERSITE AHMED DRAIA- ADRAR*. [En línea] 2021. [Citado el: 19 de noviembre de 2021.] <https://dspace.univ-adrar.edu.dz/jspui/handle/123456789/5594>.

**Shivani, Verma y Pradeep, Tomar. 2021.** *Impact of AI Technologies on Teaching, Learning, and Research in Higher Education*. Km. Mayawati Government Girls P.G. College, India : Gautam Buddha University, 2021. 1799847632.

**Simister, Nigel. 2017.** *Basic Tools for Data Collection*. United States : Intrac for civil society, 2017.

**Smith, Leslie N. 2017.** *Cyclical Learning Rates for Training Neural Networks*. Washington, D.C. : 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), DOI: 10.1109/WACV.2017.58, 2017. ISBN:978-1-5090-4823-6.

**Soltani, Mehran, y otros. 2019.** *Deep Learning-Based Channel Estimation*. s.l. : IEEE Communications Letters, vol. 23, no. 4, pp. 652-655, 2019. doi: 10.1109/LCOMM.2019.2898944.



**Téllez, Vivianne y Martínez, Duyber. 2021.** *MÉTODO AUTOMÁTICO PARA LA PREDICCIÓN DEL AVALÚO COMERCIAL DE UN INMUEBLE EN LA CIUDAD DE BOGOTÁ.* Bogotá, Colombia : Universidad Católica de Colombia, Facultad de ingeniería, 2021.

**Urbana. 2016.** Urbana.pe. *¿Cómo calcular el valor de venta de tu casa?* [En línea] 6 de setiembre de 2016. [Citado el: 16 de octubre de 2021.] <https://urbana.pe/blog/asesoria-inmobiliaria/como-calcular-el-valor-de-venta-de-tu-casa/>.

**Villada, Fernando, Muñoz, Nicolas y García-Quintero, Edwin. 2016.** *Redes Neuronales Artificiales aplicadas a la Predicción del Precio del Oro.* Medellín : Información Tecnológica, 2016. ISSN 0718-0764.

**Wang, Weijie y Lu, Yanmin. 2018.** *IOP Conference Series: Materials Science and Engineering.* Kuala Lumpur, Malaysia : IOP Publishing Ltd, 2018. 324 012049.

**Westreicher, Guillermo. 2020.** Economipedia.com. *Predicción (estadística).* [En línea] 20 de agosto de 2020. [Citado el: 2021 de octubre de 23.] <https://economipedia.com/definiciones/prediccion-estadistica.html>.

**Yang, Daniel, y otros. 2022.** *Comparison of machine learning methods for clinical data imputation among a real-world lung cancer cohort.* s.l. : MedRxiv, 2022. doi: <https://doi.org/10.1101/2022.06.12.22276306>.

**Zapata, Juan Gerardo. 2019.** *APLICACIÓN DE LAS REDES NEURONALES PARA DETERMINAR LOS PRONÓSTICOS MENSUALES DE LA INFLACIÓN PARA EL AÑO 2018.* Piura, Perú : Escuela profesional de estadística, Facultad de Ciencias, Universidad Nacional de Piura, 2019.

## ANEXOS

### ANEXO 1: Matriz de consistencia

Problema	Objetivos	Variable	Dimensiones	Indicadores	Escala de medición	Diseño metodológico
¿Cuál es la efectividad del modelo de redes neuronales en la predicción del precio de inmuebles en la ciudad de Piura?	Determinar la efectividad del modelo de redes neuronales en la predicción de un inmueble, en base a sus características.	Predicción de precios de inmuebles.	Predicción de casas en la ciudad de Piura.	Precio predicho	Intervalo	Población: fichas de registro de resultados de simulaciones. Tipo de investigación: Aplicada. Diseño de investigación: Preexperimental. Nivel de investigación: Descriptivo.
¿Cuáles son los datos que se deben obtener para los parámetros de entrada de la red neuronal?	Identificar los parámetros de entrada para el modelo de la predicción de los precios.	Modelo basado en redes neuronales	Parámetros y patrones.	Parámetros de entrada.	Intervalo	Técnicas de obtención de datos: Observación y análisis de documentos.
¿Cuál es el porcentaje de error necesario para la predicción de precios de inmuebles?	Evaluar el porcentaje de error mínimo para que la red neuronal sea efectiva.		Criterios de evaluación.	Porcentaje de error mínimo de la predicción.	Ordinal	Instrumentos para la obtención de datos: Fichas de registro y de análisis documental.
			Entrenamiento de la red.	Incremento de la precisión.	Intervalo	Técnicas para el procesamiento de datos: Hojas de cálculo
¿Cuál es el modelo adecuado para la predicción de precios?	Determinar el modelo a desarrollar para la red neuronal.		Modelo	Algoritmo		

Fuente: Elaboración Propia

## ANEXO 2: Operacionalización de variables

Variable	Definición conceptual	Dimensión	Indicador	Nivel de medición
Modelo basado en redes neuronales	Un sistema computarizado integrado por un gran conjunto de elementos simples, muy interconectados, de los cuales se procesa información debido a su estado dinámico en función de entradas y salidas (Heidari, y otros, 2019).	Entrenamiento de la red.	Incremento de la precisión.	Intervalo
		Criterios de evaluación.	Porcentaje de error mínimo de la predicción.	Ordinal
		Parámetros y patrones.	Parámetros de entrada.	Intervalo
		Modelo	Algoritmo	
Predicción de precios de inmuebles	La predicción es el anuncio de lo que se espera que puede suceder. Es un elemento importante de las ciencias, en general, pues permiten iniciar experimentos y contrastar el hecho esperado con la realidad. (Westreicher, 2020).  Los bienes inmuebles son aquellos que como su nombre lo indica son inmóviles, que no se pueden separar del terreno en el que se encuentran, pero de manera más específica de los bienes raíces, de tal modo, son terrenos o construcciones que se caracterizan como bienes físicos o jurídicamente. (Téllez, y otros, 2021)	Predicción de casas en la ciudad de Piura.	Precio predicho.	Intervalo

Fuente: Elaboración propia

ANEXO 3: Instrumentos

Ficha de registro 1 – Indicador: Aumento de la precisión.

<b>Ficha de registro - Resumen de red neuronal construida y entrenada N°</b>	
<b>Investigador: Carlo César García Palacios</b>	
<b>Tipo de prueba: test</b>	<b>Indicador: Aumento de la precisión.</b>
Datos procesados	
Tipo de variables	
Tamaño del set de entrenamiento	
Número de Epochs	
Número de neuronas	
Número de Capas	
Función de pérdida	

Ficha de registro 2 – Indicador Precio predicho

<b>Ficha de Registro: Resultado Predicción Inmueble (Casas) Red Neuronal</b>		
<b>Investigador: Carlo César García Palacios</b>		
<b>Tipo de prueba: test</b>	<b>Indicador: Precio predicho.</b>	
Simulación N°	Resultado predicción inmueble.	Valor de validación
1		
2		
3		

Ficha de registro 3 – Indicador: Parámetros de entrada, recopilación de datos

<b>Ficha de registro – Parámetros de entrada – Recopilación de información inmuebles.</b>													
<b>Investigador: Carlo César García Palacios</b>													
<b>Tipo: Recolección de información.</b>							<b>Indicador: Parámetros de entrada.</b>						
ítem	fecha	precio	M2 del terreno	M2 habitables	Condición	Código postal	vista	pisos	habitaciones	baño	cochera	piscina	terraza

Ficha de registro 4 – Indicador: Porcentaje de error mínimo de la predicción

<b>Ficha de Registro: Error mínimo de predicción de red neuronal</b>				
<b>Investigador: Carlo César García Palacios</b>				
<b>Tipo de prueba: test</b>	<b>Indicador: Porcentaje de error mínimo de la predicción</b>			
Resultado N°	Predicción red neuronal	Valor real	Error Normal	Error Absoluto
1				
2				
3				

Anexo 5: constancia de uso de datos

Piura, 12 de junio del 2022

Sra. Olga Lili Seminario Ávila

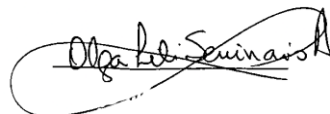
Gerente Agente Inmobiliario

Lo Tengo Inmobiliaria

Urb. Santa María del Pinar, Mz. Q, Lt. 22, calle Las Chavelas, distrito, provincia y departamento de Piura

### **CONSTANCIA**

Que el señor CARLO CÉSAR GARCÍA PALACIOS, investigador de la carrera de INGENIERÍA DE SISTEMAS de la Universidad César Vallejo, aplicó sus conocimientos e investigaciones del caso con datos de nuestra empresa, desarrollando su investigación: "Modelo basado en redes neuronales para la predicción de precios de inmuebles Piura 2021", en la cual empleó datos de diversas propiedades publicadas en venta de nuestro sitio web así como de nuestra base de datos de ventas, brindando la demostración con el código fuente del sistema y la base de datos del mismo, donde se hizo conforme la utilización correcta de los datos, culminado satisfactoriamente su periodo de investigación con nuestra empresa.



---

Olga Lili Seminario Ávila

Gerente Agente Inmobiliario N° 9176

RUC: 10027948311

## Anexo 6: Validación de instrumentos

### **CONSTANCIA DE VALIDACIÓN**

Yo, **Jaime Leandro Madrid Casariego** con DNI N° **02773132** Magister en "Ingeniería de Sistemas", de profesión Ingeniero de sistemas, desempeñándome actualmente como Docente de Ingeniería de Sistemas en la Universidad César Vallejo de Piura.

Por medio de la presente hago constar que he revisado con fines de Validación de la Ficha de registro.

Luego de hacer las observaciones pertinentes, puedo formular las siguientes apreciaciones.

Indicador de la 1: Aumento de precisión	DEFICIENTE	ACEPTABLE	BUENO	MUY BUENO	EXCELENTE
1. Claridad			x		
2. Objetividad			x		
3. Actualidad			x		
4. Organización			x		
5. Suficiencia			x		
6. Intencionalidad				x	
7. Consistencia				x	
8. Coherencia			x		
9. Metodología				x	

En señal de conformidad firmo la presente en la ciudad de Piura 07 días del mes de junio del Dos mil veintidós.

Fecha: 07 / 06 / 2022

Mg. Ing. : Jaime Leandro Madrid Casariego  
DNI : 02773132  
Especialidad : Ingeniería de Sistemas  
E-mail : jmadrid@ucv.edu.pe  
Firma :



## CONSTANCIA DE VALIDACIÓN

Yo, **Jaime Leandro Madrid Casariego** con DNI N° **02773132** Magister en "Ingeniería de Sistemas", de profesión Ingeniero de sistemas. desempeñándome actualmente como Docente de Ingeniería de Sistemas en la Universidad César Vallejo de Piura.

Por medio de la presente hago constar que he revisado con fines de Validación de la Ficha de registro.

Luego de hacer las observaciones pertinentes, puedo formular las siguientes apreciaciones.

Indicador Parámetros entrada	2: de	DEFICIEN TE	ACEPTAB LE	BUENO	MUY BUENO	EXCELEN TE
1. Claridad					x	
2. Objetividad				x		
3. Actualidad				x		
4. Organización					x	
5. Suficiencia				x		
6. Intencionalidad				x		
7. Consistencia				x		
8. Coherencia				x		
9. Metodología			x			

En señal de conformidad firmo la presente en la ciudad de Piura 07 días del mes de junio del Dos mil veintidós.

Fecha: 07 / 06 / 2022

Mg. Ing. : Jaime Leandro Madrid Casariego  
DNI : 02773132  
Especialidad : Ingeniería de Sistemas  
E-mail : jmadrid@ucv.edu.pe  
Firma :





### **CONSTANCIA DE VALIDACIÓN**

Yo, **Jaime Leandro Madrid Casariego** con DNI N° **02773132** Magister en "Ingeniería de Sistemas", de profesión Ingeniero de sistemas, desempeñándome actualmente como Docente de Ingeniería de Sistemas en la Universidad César Vallejo de Piura.

Por medio de la presente hago constar que he revisado con fines de Validación de la Ficha de registro.

Luego de hacer las observaciones pertinentes, puedo formular las siguientes apreciaciones.

<b>Indicador 3: Precio predicho</b>	<b>DEFICIENTE</b>	<b>ACEPTABLE</b>	<b>BUENO</b>	<b>MUY BUENO</b>	<b>EXCELENTE</b>
1. Claridad			x		
2. Objetividad		x			
3. Actualidad			x		
4. Organización			x		
5. Suficiencia			x		
6. Intencionalidad			x		
7. Consistencia			x		
8. Coherencia			x		
9. Metodología		x			

En señal de conformidad firmo la presente en la ciudad de Piura 07 días del mes de junio del Dos mil Veintidós.

Fecha: 07 / 06 / 2022

Mg. Ing. : Jaime Leandro Madrid Casariego  
DNI : 02773132  
Especialidad : Ingeniería de Sistemas  
E-mail : jmadrid@ucv.edu.pe  
Firma :



### **CONSTANCIA DE VALIDACIÓN**

Yo, **Jaime Leandro Madrid Casariego** con DNI N° **02773132** Magister en "Ingeniería de Sistemas", de profesión Ingeniero de sistemas. desempeñándome actualmente como Docente de Ingeniería de Sistemas en la Universidad César Vallejo de Piura.

Por medio de la presente hago constar que he revisado con fines de Validación de la Ficha de registro.

Luego de hacer las observaciones pertinentes, puedo formular las siguientes apreciaciones.

<b>Indicador</b>	<b>4:</b>	<b>DEFICIEN</b>	<b>ACEPTAB</b>	<b>BUENO</b>	<b>MUY</b>	<b>EXCELEN</b>
<b>Porcentaje de error mínimo de la predicción</b>	<b>TE</b>	<b>LE</b>	<b>TE</b>	<b>BUENO</b>	<b>BUENO</b>	<b>TE</b>
1. Claridad					x	
2. Objetividad					x	
3. Actualidad				x		
4. Organización					x	
5. Suficiencia				x		
6. Intencionalidad					x	
7. Consistencia				x		
8. Coherencia					x	
9. Metodología					x	

En señal de conformidad firmo la presente en la ciudad de Piura a los 07 días del mes de junio del Dos mil veintidós.

Fecha: 07 / 06 / 2022

Mg. Ing. : Jaime Leandro Madrid Casariego  
DNI : 02773132  
Especialidad : Ingeniería de Sistemas  
E-mail : jmadrid@ucv.edu.pe  
Firma :



### CONSTANCIA DE VALIDACIÓN

Yo, **Iván Michell Castillo Jiménez** con DNI N° **02883813** Doctor en "Tecnologías de la Información y Comunicaciones", de profesión Ingeniero Informático, desempeñándome actualmente como Docente de Ingeniería de Sistemas en la Universidad César Vallejo de Piura.

Por medio de la presente hago constar que he revisado con fines de Validación de la Ficha de registro.

Luego de hacer las observaciones pertinentes, puedo formular las siguientes apreciaciones.

Indicador de la precisión	1:	DEFICIENTE	ACEPTABLE	BUENO	MUY BUENO	EXCELENTE
1. Claridad				X		
2. Objetividad			X			
3. Actualidad			X			
4. Organización				X		
5. Suficiencia			X			
6. Intencionalidad					X	
7. Consistencia						
8. Coherencia				X		
9. Metodología			X			

En señal de conformidad firmo la presente en la ciudad de Piura 10 días del mes de junio del Dos mil veintidós.

Fecha: 10 / 06 / 2022

Mg. Ing. : Iván Michell Castillo Jiménez  
DNI : 02883813  
Especialidad : Ingeniería de Sistemas  
E-mail : icastillo@ucvvirtual.edu.pe  
Firma : 

### CONSTANCIA DE VALIDACIÓN

Yo, **Iván Michell Castillo Jiménez** con DNI N° **02883813** Doctor en "Tecnologías de la Información y Comunicaciones", de profesión Ingeniero Informático, desempeñándome actualmente como Docente de Ingeniería de Sistemas en la Universidad César Vallejo de Piura.

Por medio de la presente hago constar que he revisado con fines de Validación de la Ficha de registro.

Luego de hacer las observaciones pertinentes, puedo formular las siguientes apreciaciones.

Indicador 3: Precio predicho	DEFICIENTE	ACEPTABLE	BUENO	MUY BUENO	EXCELENTE
1. Claridad			X		
2. Objetividad		X			
3. Actualidad		X			
4. Organización			X		
5. Suficiencia		X			
6. Intencionalidad			X		
7. Consistencia		X			
8. Coherencia		X			
9. Metodología		X			

En señal de conformidad firmo la presente en la ciudad de Piura 10 días del mes de junio del Dos mil veintidós.

Fecha: 10 / 06 / 2022

Mg. Ing. : Iván Michell Castillo Jiménez  
DNI : 02883813  
Especialidad : Ingeniería de Sistemas  
E-mail : icastillo@ucvvirtual.edu.pe  
Firma : 



### **CONSTANCIA DE VALIDACIÓN**

Yo, **Iván Michell Castillo Jiménez** con DNI N° **02883813** Doctor en "Tecnologías de la Información y Comunicaciones", de profesión Ingeniero Informático. desempeñándome actualmente como Docente de Ingeniería de Sistemas en la Universidad César Vallejo de Piura.

Por medio de la presente hago constar que he revisado con fines de Validación de la Ficha de registro.

Luego de hacer las observaciones pertinentes, puedo formular las siguientes apreciaciones.

<b>Indicador Parámetros entrada</b>	<b>2: de</b>	<b>DEFICIEN TE</b>	<b>ACEPTAB LE</b>	<b>BUENO</b>	<b>MUY BUENO</b>	<b>EXCELEN TE</b>
1. Claridad				x		
2. Objetividad			x			
3. Actualidad				x		
4. Organización				x		
5. Suficiencia			x			
6. Intencionalidad					x	
7. Consistencia				x		
8. Coherencia					x	
9. Metodología			x			

En señal de conformidad firmo la presente en la ciudad de Piura 10 días del mes de junio del Dos mil veintidós.

Fecha: 10 / 06 / 2022

Mg. Ing. : Iván Michell Castillo Jiménez  
DNI : 02883813  
Especialidad : Ingeniería de Sistemas  
E-mail : icastilloj@ucvvirtual.edu.pe  
Firma :



**-CONSTANCIA DE VALIDACIÓN**

Yo, **Iván Michell Castillo Jiménez** con DNI N° **02883813** Doctor en "Tecnologías de la Información y Comunicaciones", de profesión Ingeniero Informático. desempeñándome actualmente como Docente de Ingeniería de Sistemas en la Universidad César Vallejo de Piura.

Por medio de la presente hago constar que he revisado con fines de Validación de la Ficha de registro.

Luego de hacer las observaciones pertinentes, puedo formular las siguientes apreciaciones.

Indicador Porcentaje de error mínimo de la predicción	4: DEFICIEN TE	ACEPTAB LE	BUENO	MUY BUENO	EXCELEN TE
1. Claridad		x			
2. Objetividad		x			
3. Actualidad			x		
4. Organización				x	
5. Suficiencia			x		
6. Intencionalidad			x		
7. Consistencia			x		
8. Coherencia			x		
9. Metodología		x			

En señal de conformidad firmo la presente en la ciudad de Piura 10 días del mes de junio del Dos mil veintidós.

Fecha: 10 / 06 / 2022

Mg. Ing. : Iván Michell Castillo Jiménez  
DNI : 02883813  
Especialidad : Ingeniería de Sistemas  
E-mail : icastilloj@ucvvirtual.edu.pe  
Firma :



## CONSTANCIA DE VALIDACIÓN

Yo, **Teófilo Roberto Correa Calle**, con DNI N° 02820231 Magister en "Dirección y Gestión de TI", de profesión Ingeniero en informática, desempeñándome actualmente como Docente de la escuela de ingeniería de sistemas en la Universidad César Vallejo de Piura.

Por medio de la presente hago constar que he revisado con fines de Validación de la Ficha de análisis documental.

Luego de hacer las observaciones pertinentes, puedo formular las siguientes apreciaciones.

Indicador de la precisión	1:	DEFICIENTE	ACEPTABLE	BUENO	MUY BUENO	EXCELENTE
1. Claridad					X	
2. Objetividad					X	
3. Actualidad				X		
4. Organización				X		
5. Suficiencia					X	
6. Intencionalidad					X	
7. Consistencia					X	
8. Coherencia					X	
9. Metodología				X		

Aplicabilidad: El instrumento puede ser aplicado (X)

El instrumento debe ser mejorado ( )

Observaciones:

---

En señal de conformidad firmo la presente en la ciudad de Piura a los 08 días del mes de junio del Dos mil veinte.

Fecha: 21/11/2021

Mg. Ing. : Teófilo Roberto Correa Calle  
DNI : 02820231  
Especialidad : Ingeniería en Informática  
E-mail : terococa@gmail.com  
Firma :



TEÓFILO ROBERTO  
CORREA CALLE

## CONSTANCIA DE VALIDACIÓN

Yo, Teófilo Roberto Correa Calle, con DNI N° 02820231 Magister en "Dirección y Gestión de TI", de profesión Ingeniero en informática. desempeñándome actualmente como Docente de la escuela de ingeniería de sistemas en la Universidad César Vallejo de Piura.

Por medio de la presente hago constar que he revisado con fines de Validación de la Ficha de análisis documental.

Luego de hacer las observaciones pertinentes, puedo formular las siguientes apreciaciones.

Indicador Parámetros entrada	2: de	DEFICIENTE	ACEPTABLE	BUENO	MUY BUENO	EXCELENTE
1. Claridad					X	
2. Objetividad					X	
3. Actualidad					X	
4. Organización				X		
5. Suficiencia				X		
6. Intencionalidad					X	
7. Consistencia					X	
8. Coherencia					X	
9. Metodología				X		

Aplicabilidad: El instrumento puede ser aplicado (X)

El instrumento debe ser mejorado ()

Observaciones:

---

En señal de conformidad firmo la presente en la ciudad de Piura a los 08 días del mes de junio del Dos mil veinte.

Fecha: 21 / 11 / 2021

Mg. Ing. : Teófilo Roberto Correa Calle  
DNI : 02820231  
Especialidad : Ingeniería en Informática  
E-mail : terococa@gmail.com



TEÓFILO ROBERTO  
CORREA CALLE



### **CONSTANCIA DE VALIDACION**

Yo, Teófilo Roberto Correa Calle, con DNI N° 02820231 Magister en "Dirección y Gestión de TI", de profesión Ingeniero en informática. desempeñándome actualmente como Docente de la escuela de ingeniería de sistemas en la Universidad César Vallejo de Piura.

Por medio de la presente hago constar que he revisado con fines de Validación de la Ficha de análisis documental.

Luego de hacer las observaciones pertinentes, puedo formular las siguientes apreciaciones.

Indicador 3:	DEFICIENTE	ACEPTABLE	BUENO	MUY BUENO	EXCELENTE
1. Claridad				X	
2. Objetividad				X	
3. Actualidad			X		
4. Organización			X		
5. Suficiencia			X		
6. Intencionalidad			X		
7. Consistencia			X		
8. Coherencia				X	
9. Metodología				X	

Aplicabilidad: El instrumento puede ser aplicado ( )

El instrumento debe ser mejorado ( )

Observaciones:

En señal de conformidad firmo la presente en la ciudad de Piura a los 08 días del mes de junio del Dos mil veinte.

Fecha: 21 / 11 / 2021

Mg. Ing. : Teófilo Roberto Correa Calle  
DNI : 02820231  
Especialidad : Ingeniería en Informática  
E-mail : terococa@gmail.com



TEOFILO ROBERTO CORREA CALLE

### **CONSTANCIA DE VALIDACIÓN**

Yo, Teófilo Roberto Correa Calle, con DNI N° 02820231 Magister en "Dirección y Gestión de TI", de profesión Ingeniero en informática. desempeñándome actualmente como Docente de la escuela de ingeniería de sistemas en la Universidad César Vallejo de Piura.

Por medio de la presente hago constar que he revisado con fines de Validación de la Ficha de análisis documental.

Luego de hacer las observaciones pertinentes, puedo formular las siguientes apreciaciones.

Indicador 4:	DEFICIENTE	ACEPTABLE	BUENO	MUY BUENO	EXCELENTE
1. Claridad				X	
2. Objetividad				X	
3. Actualidad			X		
4. Organización			X		
5. Suficiencia			X		
6. Intencionalidad			X		
7. Consistencia			X		
8. Coherencia				X	
9. Metodología				X	

Aplicabilidad: El instrumento puede ser aplicado ( )

El instrumento debe ser mejorado (X)

Observaciones:

En señal de conformidad firmo la presente en la ciudad de Piura a los 08 días del mes de junio del Dos mil veinte.

Fecha: 21 / 11 / 2021

Mg. Ing. : Teófilo Roberto Correa Calle  
DNI : 02820231  
Especialidad : Ingeniería en Informática  
E-mail : terococa@gmail.com



TEÓFILO ROBERTO  
CORREA CALLE

## Anexo 7: Certificado de traducción de resumen – abstract

Piura, 21 de junio del 2022

### CERTIFICADO DE TRADUCCIÓN

Yo, HENRY WYATT LEUE MAZA, con DNI 78290268, me encuentro en capacidad para traducir del idioma español al idioma inglés, y certifico que la traducción del texto:

“The objective of this investigation was to determine the effectiveness of a neural network model in predicting real estate prices in Piura 2021. The focus of this research is quantitative, because the model was analyzed using descriptive and regression statistics; therefore, the type of research is applied; the research design is pre-experimental and explanatory. Among the results, it was found that the sequential model was the one implemented because it is based on a function of several inputs and a single output. A processed training set of 21,000 properties was obtained, including sales made and listed for sale. Information was also obtained showing that the neural network had 3 simulations, where it yielded a variance score of 0.8 in the third simulation and an effectiveness of 25% as a minimum result, using seven input variables for the model and for its subsequent validation. In conclusion, it was determined that the ANN model is effective not only because of the correct configurations made after each simulation, but also because of the use of sklearn metrics for regression and the most optimal model selection, which allowed the model to be evaluated based on its accuracy during training.”

es verdadero y exacto en la medida de mis posibilidades.



FIRMA

**HENRY WYATT LEUE MAZA**

**78290268**

Docente del Instituto Cultural Peruano Norteamericano

**HLEUE@ICPNA-RG.EDU.PE**

## ANEXO 8: Metodología usada para implementar Red Neuronal

### A) Definición de la red neuronal

- Selección, limpieza y escalamiento del set de entrenamiento

```
import tensorflow as tf
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

##Importando Datos
house_df = pd.read_csv("Reco_inmuebles.csv")

##LIMPIEZA DE DATOS
selected_features = ['bedrooms', 'bathrooms', 'sqft_living', 'sqft_lot', 'floors']

X = house_df[selected_features]
y = house_df['price']

from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
X_scaled = scaler.fit_transform(X)

#Normalizando output
y = y.values.reshape(-1,1)
y_scaled = scaler.fit_transform(y)
```

- Definición del modelo

```
compiler flags.
Model: "sequential"

Layer (type)                 Output Shape              Param #
-----
dense (Dense)                 (None, 19)                114
dense_1 (Dense)               (None, 19)                380
dense_2 (Dense)               (None, 19)                380
dense_3 (Dense)               (None, 1)                 20
-----
Total params: 894
Trainable params: 894
```

### B) Entrenamiento de la RNA

- Uso de tipo de aprendizaje (supervisado)

```
##ENTRENAMIENTO
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X_scaled, y_scaled, test_size = 0.3, random_state=101)

##Definiendo modelo
model = tf.keras.Sequential([
    tf.keras.layers.Dense(19),
    tf.keras.layers.Dense(19),
    tf.keras.layers.Dense(1)
```

- Entrenamiento basado en Epochs

```
val_loss: 9.0097e-04
Epoch 8/400
95/95 [=====] - 0s 2ms/step - loss: 0.0011 -
val_loss: 9.3194e-04
Epoch 9/400
95/95 [=====] - 0s 2ms/step - loss: 0.0011 -
val_loss: 8.8636e-04
Epoch 10/400
95/95 [=====] - 0s 2ms/step - loss: 0.0011 -
val_loss: 9.0332e-04
Epoch 11/400
95/95 [=====] - 0s 2ms/step - loss: 0.0011 -
val_loss: 8.8924e-04
Epoch 12/400
1/95 [.....] - ETA: 0s - loss: 8.5377e-04
```

```
In [4]: losses = pd.DataFrame(model.history.history)
...: losses
Out[4]:
```

	loss	val_loss
0	0.001459	0.000961
1	0.001186	0.000935
2	0.001132	0.000920
3	0.001103	0.000905
4	0.001086	0.000923
..	...	...
395	0.000918	0.000841
396	0.000915	0.000839
397	0.000920	0.000835
398	0.000914	0.000850
399	0.000912	0.000860

C) Validación del modelo (Utilización)

- Uso del modelo con función model.predict

```
#Predicción
#Definir inmueble para la predicción con las características / inputs
# 'bedrooms','bathrooms','sqft_living','sqft_lot','floors'

#https://inmobiliarialotengo.com/propiedades/103-Casa-Piura-USD280-000
X_test_2 = np.array([[8, 5, 420, 435, 3]])
#
scaler_2 = MinMaxScaler()
X_test_scaled_2 = scaler_2.fit_transform(X_test_2)
#Haciendo predicción
y_predict_2 = model.predict(X_test_scaled_2)
#Revirtiendo Escalado para apreciar el precio correctamente escalado
y_predict_2 = scaler.inverse_transform(y_predict_2)
```