

Aplicaciones de las Redes Neuronales al mercado inmobiliario en Córdoba

Autor: Alberto Cano Valenzuela
Tutor del Trabajo: José Rafael Caro Barrera

Resumen. Este estudio se centra en desarrollar un modelo de predicción de precios de viviendas en el mercado inmobiliario de Córdoba utilizando redes neuronales. Dada la limitada disponibilidad de datos en línea, se llevó a cabo una simulación de datos propios, considerando características específicas como ubicación y tamaño de las viviendas. El proceso de desarrollo del modelo implicó entrenar una red neuronal con estos datos simulados, aplicando técnicas de preprocesamiento para garantizar la calidad y coherencia de la información. Los resultados obtenidos fueron prometedores, evidenciando la eficacia de las redes neuronales en la predicción de precios de venta. Este estudio destaca el potencial de las redes neuronales para mejorar la toma de decisiones en el mercado inmobiliario, ofreciendo a los profesionales herramientas efectivas para la estimación de precios e identificación de oportunidades de inversión.

Palabras clave. Mercado Inmobiliario, Redes Neuronales, Simulación, Predicción de precios.

Abstract. This study explores the development of a predictive model for housing prices in the real estate market of Córdoba using neural networks. Due to limited online data availability, a simulation of proprietary data was conducted, considering specific features such as location and size of the residences. The model development process involved training a neural network with these simulated data and applying preprocessing techniques to ensure data quality. The obtained results were promising, demonstrating the effectiveness of neural networks in predicting sales prices. This study underscores the potential of neural networks to enhance decision-making in the real estate market, providing professionals with powerful tools for price estimation and investment opportunity identification.

Keywords. Real Estate Market, Neural Networks, Data Simulation, Price Prediction.

1 Introducción

1.1 Contexto y justificación del estudio

El mercado inmobiliario es uno de los pilares de la economía de cualquier ciudad, y en Córdoba, esto no es una excepción.

En los últimos años, este mercado ha experimentado importantes cambios, desde la crisis financiera provocada por el colapso de la burbuja inmobiliaria en 2008 hasta los cambios de relación con las viviendas causados por la pandemia del COVID-19, en la que los ciudadanos tienen más en cuenta la búsqueda de espacios abiertos, con luz natural y en los que poder teletrabajar. (Vaquero, 2022)

Durante la elaboración de este Trabajo de Fin de Grado se explorarán las aplicaciones de las redes neuronales en el mercado inmobiliario de Córdoba. Esta ciudad, conocida por su rica historia y patrimonio cultural, es también un importante centro económico y turístico de la región de Andalucía.

En los últimos años, el mercado inmobiliario de Córdoba ha experimentado un crecimiento significativo gracias a la llegada de nuevos residentes y turistas y a la creciente demanda de propiedades en la ciudad, tanto como que,

en 2021, la ciudad de Córdoba comercializó la mayor cantidad de pisos desde 2008. (¿Cuánto ha subido el precio de la vivienda en Córdoba?, 2023)

En este contexto, el uso de tecnologías innovadoras como las redes neuronales puede ser de gran ayuda para entender mejor el mercado inmobiliario de nuestra ciudad, pudiéndose utilizar para mejorar la eficiencia y la rentabilidad de los agentes inmobiliarios, así como para proporcionar a los clientes información más precisa y relevante sobre las propiedades disponibles en el mercado.

1.2 Objetivos del trabajo

El objetivo principal de este Trabajo de Fin de Grado es profundizar en el uso de las redes neuronales para la predicción de precios de bienes raíces en el mercado inmobiliario, en concreto, para la elaboración de este trabajo, nos centraremos en el mercado inmobiliario de la ciudad de Córdoba. Para lograr este objetivo, se llevará a cabo un estudio detallado sobre la aplicación de estas técnicas de aprendizaje automático en la evaluación de los precios de las viviendas en el mercado inmobiliario local.

Con el fin de comprender en mayor profundidad los factores que influyen en el mercado inmobiliario de Córdoba y su relación con las redes neuronales, se llevará a cabo un análisis de las características de las viviendas y de otros factores que pueden afectar su valor. A través de este análisis, se buscará identificar las variables más relevantes en la determinación del precio de una vivienda en la ciudad, lo que permitirá obtener una comprensión más precisa del mercado inmobiliario local.

Otro objetivo importante de este trabajo será evaluar la precisión y la eficacia de las redes neuronales en la predicción de los precios de las viviendas en Córdoba. Se analizará el rendimiento de dichas redes en la predicción de los precios de las viviendas.

Por último, se buscará identificar las posibles limitaciones y desafíos en la aplicación de las redes neuronales en el mercado inmobiliario de Córdoba como puede ser algo tan primordial como la imposibilidad de obtener datos para entrenar la red neuronal.

1.3 Metodología empleada y fuente de datos

En el presente Trabajo de Fin de Grado, se utiliza una metodología específica junto con fuentes de datos seleccionadas para abordar el objetivo de investigación planteado. El enfoque metodológico adoptado se basó en la recopilación manual de datos de viviendas del portal inmobiliario Fotocasa. Estos datos fueron posteriormente procesados utilizando el entorno de desarrollo PyCharm y un código de Python desarrollado de manera personalizada para abordar los objetivos de este trabajo.

La elección de Fotocasa como fuente de datos se debió a su amplia oferta de propiedades inmobiliarias y a la disponibilidad de información detallada sobre viviendas diferenciadas por barrios. Los datos utilizados en este estudio fueron recolectados manualmente mediante la exploración y búsqueda en la página web de Fotocasa, específicamente enfocándose en los datos disponibles para diferentes barrios de interés.

Una vez recopilados los datos de las viviendas de Fotocasa, se procedió a su procesamiento utilizando el entorno de desarrollo integrado. Se desarrolló un código personalizado en Python con el propósito de generar datos simulados. Mediante este código, se crearon variables de ajuste utilizando los datos que se extrajeron previamente, como el precio medio por metro cuadrado o la superficie en metros cuadrados que se repetían con mayor frecuencia en las ventas de viviendas, entre otros datos de interés. Utilizando estos valores como referencia, se generaron datos simulados que se ajustaban en torno a dicha media, lo que ha permitido obtener un conjunto amplio de datos para el estudio.

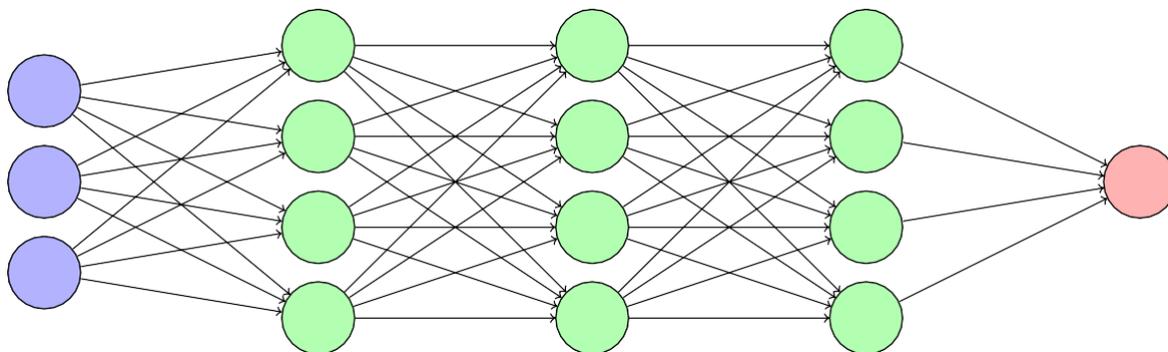
2 Marco teórico

2.1 Conceptos básicos de las redes neuronales y sus aplicaciones

Una red neuronal es un modelo computacional inspirado en la estructura y funcionamiento del cerebro humano, diseñado para aprender a través de la experiencia y procesar información de manera similar a como lo hace nuestro cerebro. (Neural Networks, s.f.)

Una red neuronal está compuesta por un conjunto de nodos interconectados, llamados neuronas artificiales o unidades de procesamiento, que reciben señales de entrada, realizan cálculos y producen señales de salida. Estas neuronas están organizadas en capas, y las señales se propagan a través de la red de manera que cada capa procesa información cada vez más abstracta y compleja. (Di Pietro, 2021)

Figura 1: Estructura de una red neuronal



Fuente: Elaboración propia

El proceso de entrenamiento de una red neuronal implica proporcionarle un conjunto de datos de entrada y una salida esperada, y ajustar los pesos de las conexiones entre las neuronas para que la red produzca la salida deseada a partir de la entrada dada. Una vez que una red neuronal ha sido entrenada, puede ser utilizada para predecir salidas para nuevas entradas que no se hayan visto anteriormente. (Mesquita, 2021)

Las redes neuronales se utilizan en una amplia variedad de aplicaciones, incluyendo reconocimiento de imágenes, procesamiento del lenguaje natural, análisis de datos, robótica o control automático, entre otras muchas.

Hay muchos tipos de redes neuronales, cada una con su propia estructura y función específicas. Algunos de los tipos más comunes de redes neuronales son los siguientes:

- Redes neuronales *feedforward* (FNN), también conocidas como redes neuronales de propagación hacia adelante, son un tipo de modelo de aprendizaje profundo en el que la información fluye en una sola dirección, de la entrada a la salida.

En estas redes, las neuronas se organizan en capas: una capa de entrada, una o más capas ocultas y una capa de salida. Cada neurona en una capa se conecta a todas las neuronas de la capa siguiente.

Las entradas se propagan a través de la red de manera secuencial, donde cada capa de neuronas procesa los datos que recibe y los envía a la siguiente capa hasta llegar a la capa de salida. Cada neurona en una capa procesa la información que recibe mediante una función de activación, y las conexiones entre las neuronas están ponderadas, lo que significa que cada conexión tiene un peso asociado que modifica la importancia de esa conexión en la salida de la neurona.

Las redes neuronales *feedforward* son muy utilizadas en tareas de clasificación y regresión, y pueden ser entrenadas mediante algoritmos de optimización como el descenso del gradiente estocástico. (Turing, 2022)

- Las redes neuronales recurrentes (RNN) son utilizadas en el procesamiento de secuencias de datos.

Este tipo de redes neuronales son especialmente adecuadas para el procesamiento de datos secuenciales que cambian con el tiempo, como el habla, el lenguaje natural, las series temporales y las señales de audio. Estas redes están compuestas por neuronas organizadas en capas que se conectan entre sí y se retroalimentan a través del tiempo, lo que les permite mantener una “memoria” a corto plazo de los datos de entrada anteriores y utilizarla para informar la salida actual.

La retroalimentación de este tipo de redes puede hacer que los gradientes se vuelvan inestables y dificultar su entrenamiento. Sin embargo, se han desarrollado variantes de redes neuronales recurrentes, como las redes neuronales de memoria a largo plazo y las redes neuronales de puerta de enlace, para abordar este problema y mejorar su rendimiento.

Las redes neuronales recurrentes se utilizan en diversas aplicaciones, como el reconocimiento de voz, la traducción automática, la generación de texto, el análisis de sentimientos y la predicción de series temporales. Estas redes tienen la capacidad de aprender de patrones a lo largo del tiempo y procesar secuencias de longitud variable, lo que las hace muy versátiles en el campo del aprendizaje automático. (Introducción a las Redes Neuronales Recurrentes, 2019)

- Redes neuronales convolucionales (CNN) son un tipo de algoritmo de aprendizaje profundo utilizado en el procesamiento de imágenes y señales de audio.

Las CNN se basan en el funcionamiento de las células nerviosas del cerebro humano, que procesan la información a través de la interconexión de neuronas. Al igual que el cerebro humano, las CNN utilizan capas de neuronas interconectadas para procesar la información.

Las capas de neuronas de una CNN se dividen en dos tipos principales: capas convolucionales y capas de agrupación. Las capas convolucionales aplican un filtro a la imagen de entrada para extraer características importantes, como bordes y texturas, mientras que las capas de agrupación reducen la dimensionalidad de las características para acelerar el proceso de aprendizaje.

Una vez que se han extraído las características de la imagen, se alimentan a través de una serie de capas de neuronas completamente conectadas que clasifican la imagen en una de varias categorías predefinidas.

Las CNN se utilizan en una amplia variedad de aplicaciones de procesamiento de imágenes y señales de audio, incluyendo reconocimiento de imágenes, detección de objetos, clasificación de audio y análisis de sentimientos en el procesamiento del lenguaje natural. (Saha, 2018)

- Redes neuronales *autoencoder* son un tipo de algoritmo de aprendizaje profundo que se utiliza para la reducción de la dimensionalidad, la compresión de datos y la reconstrucción de señales.

El objetivo principal de las redes neuronales *autoencoder* es aprender una representación comprimida de los datos de entrada que contenga solo la información más importante. El *autoencoder* consiste en dos partes principales: un codificador y un decodificador.

El codificador recibe una entrada de alta dimensionalidad y la reduce a una representación de baja dimensionalidad que conserva la información más relevante. El decodificador toma esta representación de baja dimensionalidad y la utiliza para reconstruir la entrada original con la mayor precisión posible.

Durante el entrenamiento, el *autoencoder* se ajusta para minimizar la diferencia entre la entrada original y la salida reconstruida. Esto obliga al *autoencoder* a aprender una representación de baja dimensionalidad que contiene solo la información más relevante para la entrada.

Los *autoencoder* se utilizan en una amplia variedad de aplicaciones, como la compresión de imágenes, la eliminación de ruido en señales de audio y la reducción de la dimensionalidad de datos para la visualización. También se utilizan a menudo como componentes de sistemas más grandes, como sistemas de recomendación y sistemas de reconocimiento de patrones. (Paredes, 2020)

Cada tipo de red neuronal tiene sus propias fortalezas y debilidades, la elección de la red más adecuada dependerá de las características específicas de los datos y la tarea a la que se aplique. En general, las redes neuronales *feedforward* son efectivas en tareas de clasificación y regresión de datos no secuenciales, mientras que las redes neuronales recurrentes son ideales para procesar datos secuenciales o dinámicos. Por su parte, las redes neuronales convolucionales son especialmente útiles en tareas de procesamiento de imágenes y visión por computadora. (Kaushik, 2021)

Sin embargo, es importante tener en cuenta que no existe una única “mejor” red neuronal para una tarea determinada. En lugar de eso, la elección de la red neuronal más adecuada dependerá de diversos factores, como la cantidad y la calidad de los datos disponibles, el tiempo disponible para el entrenamiento del modelo y la disponibilidad de recursos computacionales.

No debemos olvidar que el campo del aprendizaje automático se encuentra en constante evolución, y se están desarrollando constantemente nuevas técnicas y arquitecturas de red neuronal. Es importante seguir investigando

y evaluando estas nuevas técnicas para estar al tanto de las últimas innovaciones en el campo del aprendizaje automático.

La elección de la red neuronal adecuada es crucial para el éxito de cualquier proyecto de aprendizaje automático. Esta elección dependerá de las características específicas de los datos y la tarea, así como de la disponibilidad de recursos computacionales. Es importante seguir investigando y evaluando las nuevas técnicas y arquitecturas de red neuronal a medida que se desarrollan para mantenerse al día con las últimas innovaciones en el campo del aprendizaje automático.

2.2 Aplicaciones de redes neuronales utilizados en la predicción de valores inmobiliarios

Las redes neuronales son una herramienta útil y poderosa para la predicción de valores inmobiliarios debido a su capacidad para analizar grandes cantidades de datos y reconocer patrones complejos lo que permite a los agentes y tasadores inmobiliarios determinar el valor de una propiedad de manera más precisa y eficiente. Las redes neuronales pueden analizar datos sobre características específicas de la propiedad, como el tamaño, la ubicación, la antigüedad y el estado en el que se encuentra, así como datos de ventas recientes de propiedades similares en el área. Al entrenar una red neuronal con estos datos, se puede desarrollar un modelo que prediga el valor de una propiedad en función de sus características. (Kalliola, Kapočiūtė-Dzikienė, y Damaševičius, 2021)

Además de la predicción de valores inmobiliarios, las redes neuronales también pueden utilizarse en una variedad de otras aplicaciones relacionadas con bienes raíces, como la identificación de oportunidades de inversión, el análisis de mercado o la evaluación de riesgos.

Al utilizar redes neuronales para la identificación de oportunidades de inversión, se pueden analizar datos históricos de ventas de propiedades en un área determinada para identificar patrones y tendencias que puedan indicar un potencial de inversión rentable. Por ejemplo, una red neuronal puede identificar una tendencia en el aumento de los precios de las propiedades en un área determinada debido a factores como el desarrollo de nuevas infraestructuras o un aumento en la demanda de propiedades. (Choy y Ho, 2023)

Del mismo modo, las redes neuronales pueden utilizarse en el análisis de mercado para predecir tendencias futuras y tomar decisiones informadas sobre la compra y venta de propiedades. Por ejemplo, una red neuronal puede analizar datos sobre la economía local, las tasas de interés y la oferta y demanda de propiedades para predecir la dirección del mercado inmobiliario y determinar la mejor estrategia de inversión. En la evaluación de riesgos, las redes neuronales pueden utilizarse para analizar datos históricos sobre factores como la tasa de delincuencia, las inundaciones y los terremotos para determinar el nivel de riesgo asociado con una propiedad. Al entrenar una red neuronal con datos históricos de desastres naturales y crímenes, se puede desarrollar un modelo que prediga el nivel de riesgo asociado con una propiedad en función de su ubicación y características. Dentro de la evaluación de riesgos, las redes neuronales pueden analizar patrones de comportamiento en transacciones inmobiliarias para identificar posibles fraudes. (Chaipimonplina y Sin-Ampola, 2019; RB y KR, 2021; Walczak, 2021)

3 Generación de datos simulados con Python

3.1 Problema de la falta de datos en el mercado inmobiliario

El problema del análisis de datos en el contexto del mercado inmobiliario radica en la obtención de datos relevantes y confiables para alimentar los modelos de análisis y predicción. Esta etapa es fundamental, ya que los resultados y la calidad de los modelos dependen en gran medida de la calidad de los datos utilizados. (Ruiz, 2022).

La escasez de datos disponibles y las limitaciones de las fuentes existentes son desafíos significativos en el contexto del mercado inmobiliario. Estos problemas dificultan la aplicación efectiva de las redes neuronales y otras técnicas de análisis de datos en la predicción y valoración de propiedades inmobiliarias.

En primer lugar, la falta de datos realistas y completos puede deberse a diversas razones. Por un lado, la recopilación de datos inmobiliarios a gran escala puede ser costosa y requerir mucho tiempo. Las fuentes tradicionales, como las bases de datos públicas y los registros oficiales, pueden tener limitaciones en cuanto a la cantidad y calidad de los datos disponibles. Además, los datos recopilados pueden no reflejar adecuadamente la diversidad y la dinámica del mercado inmobiliario, lo que dificulta la generación de modelos precisos y confiables.

Además, las fuentes de datos existentes pueden presentar limitaciones en términos de su disponibilidad y acceso. Algunos datos relevantes, como los precios de venta reales, pueden ser confidenciales y no estar disponibles públicamente. Esto restringe aún más el conjunto de datos que se puede utilizar para entrenar y evaluar modelos de redes neuronales en el mercado inmobiliario.

La falta de datos puede tener un impacto significativo en la precisión y confianza de las predicciones y valoraciones inmobiliarias. Sin una cantidad suficiente de datos representativos y relevantes, los modelos de redes neuronales pueden tener dificultades para capturar las complejidades y las tendencias del mercado. Esto puede llevar a estimaciones inexactas de los precios de las propiedades, dificultando la toma de decisiones informadas por parte de los profesionales del sector inmobiliario.

Ante estas limitaciones, la generación de datos propios se presenta como una solución, sin embargo, es importante tener en cuenta que la generación de datos propios no está exenta de desafíos y limitaciones. En primer lugar, generar datos reales a gran escala puede ser extremadamente costoso tanto en términos de tiempo como de recursos financieros. Recopilar datos de propiedades inmobiliarias de manera exhaustiva implica llevar a cabo inspecciones físicas, acceder a registros públicos y recabar información detallada sobre transacciones y características de las propiedades. Esto puede resultar una tarea laboriosa y costosa, especialmente si se pretende cubrir una amplia gama de propiedades y ubicaciones geográficas. (Barrios Arce, 2023)

Por otro lado, la generación de datos ficticios también plantea desafíos. Aunque es posible generar datos sintéticos utilizando técnicas de modelado y simulación, existe el riesgo de que los datos generados no sean completamente realistas. Los modelos de generación de datos pueden no capturar todas las complejidades y variaciones presentes en el mercado inmobiliario real, lo que puede afectar la calidad y la utilidad de los datos generados. Es esencial validar y evaluar cuidadosamente los datos generados para garantizar que sean representativos y se ajusten a los patrones y comportamientos del mercado inmobiliario.

3.2 Análisis de datos generados

3.2.1 Análisis de datos generados

El conjunto de datos generado consta de información simulada de ventas de viviendas, que fue generada a partir de variables ajustadas utilizando un código personalizado en Python. Este conjunto de datos ampliado se creó con el propósito de enriquecer y diversificar los datos de muestra obtenidos manualmente.

El conjunto de datos incluye atributos clave como son la zona en la que se encuentra la vivienda, la superficie en metros cuadrados, el número de habitaciones, el número de baños, si la vivienda tiene terraza, si la vivienda tiene ascensor, si la vivienda ofrece parking, si la vivienda está previamente amueblada y su valor. Estos atributos se generaron en torno a valores de referencia extraídos de los datos originales, como la media del precio por metro cuadrado y la superficie que se repetía con mayor frecuencia en las ventas de viviendas.

Se aplicaron técnicas de generación de datos sintéticos para asegurar que los nuevos datos fueran representativos y coherentes con la realidad del mercado inmobiliario. Asimismo, se tuvo en cuenta la distribución y variabilidad adecuada de cada atributo para mantener la validez de los resultados obtenidos.

Además, se utilizaron variables de ajuste para garantizar que los datos generados fueran más realistas y coherentes con la realidad del mercado inmobiliario. Para lograr esto, se utilizó un factor de ajuste del 80% en lugar del 100% para el precio medio por metro cuadrado obtenido de Fotocasa.

Este ajuste del 80% se aplicó considerando las variables adicionales, como “tiene terraza”, “tiene ascensor”, “tiene parking”, entre otras. Se observó que estas variables influyen en el precio de una vivienda, aumentándolo en muchos casos. Al ajustar el precio generado en función de estas variables, se logró que los datos simulados reflejaran mejor las condiciones del mercado inmobiliario.

Si se hubiera utilizado el 100% como factor de ajuste, el precio generado para viviendas con características adicionales podría haber resultado demasiado elevado en comparación con la realidad. Al utilizar el 80% como factor de ajuste, se logró que el precio generado fuera más acorde con la media ajustada de Fotocasa y permitió obtener valores más realistas para las viviendas generadas.

Este enfoque de ajuste del 80% en el precio medio por metro cuadrado y la consideración de variables adicionales permitió que los datos simulados reflejaran de manera más precisa las fluctuaciones y características del mercado

inmobiliario. Esto a su vez proporciona una base más sólida para el análisis y la interpretación de los resultados obtenidos en el estudio

Es importante destacar que la probabilidad de generar viviendas con características adicionales, como si la vivienda tiene terraza, ascensor, parking, o si está amueblado, se basó en la demanda real que experimenta el mercado inmobiliario en la ciudad de Córdoba.

Para asegurar la representatividad de los datos simulados, se tuvo en cuenta la información sobre la demanda y las preferencias de los compradores de viviendas en Córdoba. Esto permitió ajustar la generación de viviendas con las características mencionadas según la frecuencia observada en el mercado real.

El conjunto de datos generado se presenta en formato tabular, con cada instancia representando una vivienda simulada. Se proporciona información detallada y estructurada que permite su análisis y posterior utilización en los distintos modelos y algoritmos empleados en este estudio.

Cabe destacar que este conjunto de datos simulados se utiliza con fines puramente investigativos y no corresponde a datos reales de transacciones inmobiliarias. Sin embargo, su generación cuidadosa y el ajuste a valores de referencia permiten obtener resultados que pueden ser extrapolados y utilizados para análisis y evaluaciones relevantes en el contexto de este trabajo.

* El código desarrollado para este proyecto está disponible para su libre uso para cualquier persona que pueda encontrarlo útil a través del siguiente enlace: <https://github.com/xKnito/Generacion-datos-simulados-TFG>

3.2.2 Análisis exploratorio y visualización de los datos generados

Tabla 1: Media - Características de las viviendas

MEDIA	
Media Metros Cuadrados	110,341
Media Habitaciones	3,321
Media Baños	2,009
Media Terraza (%)	72,4
Media Ascensor (%)	29
Media Parking (%)	24,6
Media Amueblado (%)	48,9
Media Valor (€)	155247

Fuente: Elaboración propia

El promedio es una medida estadística que representa el valor central que resume la tendencia central del conjunto de datos. Teniendo esto en cuenta, en promedio, las viviendas en el conjunto de datos tienen un tamaño de alrededor de 110,341 metros cuadrados.

El promedio de habitaciones por vivienda es de aproximadamente 3,321, lo que sugiere que en promedio las viviendas tienen 3 habitaciones.

En cuanto a los baños, el promedio es de aproximadamente 2,009, lo que indica que en promedio las viviendas cuentan con 2 baños.

El promedio de la variable "Terraza" indica que alrededor del 72,4% de las viviendas en el conjunto de datos tienen terraza.

En cuanto al "Ascensor", indica que alrededor del 29% de las viviendas en promedio tienen ascensor.

Para la variable "Parking", indica que aproximadamente el 24,6% de las viviendas en promedio cuentan con espacio de estacionamiento.

El promedio de la variable "Amueblado" indica que alrededor del 48,9% de las viviendas están amuebladas.

Finalmente, el promedio del valor de las viviendas es de 155247 euros.

Tabla 2: Desviación estándar - Características de las viviendas

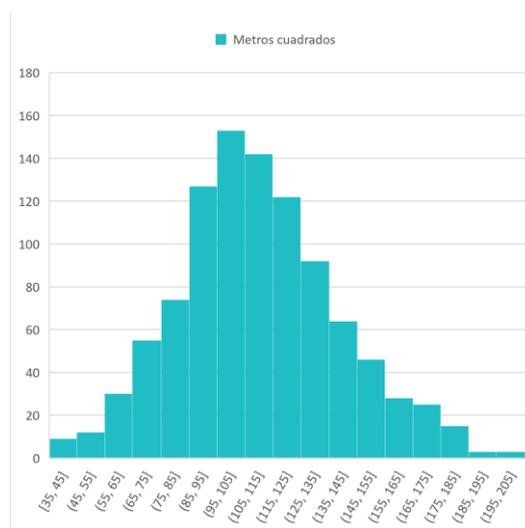
DESVIACIÓN ESTÁNDAR	
Desviación Estándar Metros Cuadrados	28,930
Desviación Estándar Habitaciones	0,618
Desviación Estándar Baños	0,386
Desviación Estándar Terraza	0,447
Desviación Estándar Ascensor	0,454
Desviación Estándar Parking	0,431
Desviación Estándar Amueblado	0,500
Desviación Estándar Valor	54410,005

Fuente: Elaboración propia

La desviación estándar de la variable “Metros Cuadrados” en el conjunto de datos es aproximadamente 28,930, lo que indica que hay una variabilidad significativa en el tamaño de las viviendas. Para el número de habitaciones, la desviación estándar es de 0,618, lo que implica que la mayoría de las viviendas tienen una cantidad de habitaciones cercana al promedio. En cuanto a los baños, la desviación estándar es de 0,386, lo que indica que la mayoría de las viviendas tienen un número de baños cercano al promedio.

Para las variables “Terraza”, “Ascensor”, “Parking” y “Amueblado”, las desviaciones estándar son de 0,447, 0,454, 0,431 y 0,500 respectivamente. Esto sugiere que hay cierta variabilidad en la presencia de estas características en las viviendas, aunque en general, la mayoría de las viviendas tienden a tener valores cercanos al promedio. La desviación estándar de la variable “Valor” es de aproximadamente 54410,005, lo que indica una alta variabilidad en los precios de las viviendas en el conjunto de datos. Considerando que la media de los precios es de 155247, se puede inferir que la mayoría de los datos (alrededor del 68%) se encuentran dentro del rango entre 100837 y 209657 ($155247 \pm 54410,005$). Esta amplia franja de valores refleja la diversidad de precios que se observa en el conjunto de datos de viviendas.

Figura 2: Distribución de los metros cuadrados de las viviendas



Fuente: Elaboración propia

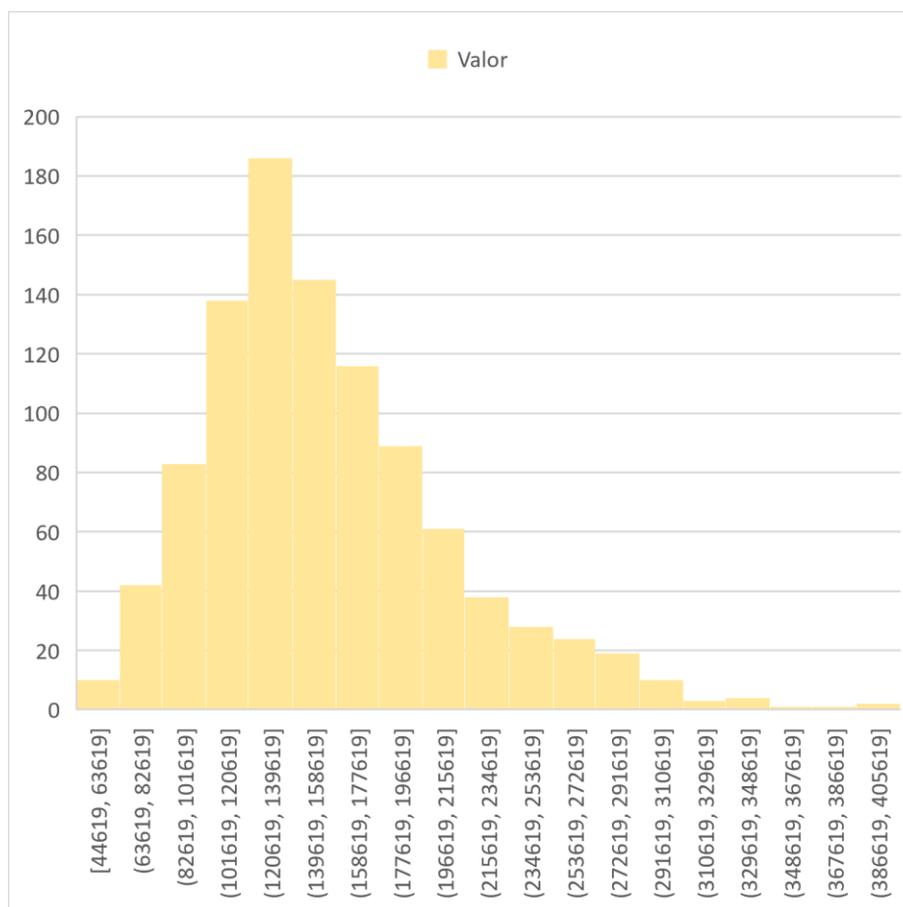
De forma visual, al realizar un histograma dividido por rangos de metros cuadrados según los datos disponibles, se puede observar la distribución de las viviendas en función de su tamaño. Vemos que la mayoría de las viviendas se encuentran en un intervalo específico. Dentro de este rango, el valor que se repite con mayor frecuencia es 95-105 metros cuadrados, indicando una concentración significativa en esa superficie.

A medida que nos alejamos de este rango hacia valores más altos o más bajos en cuanto a metros cuadrados se refiere, la cantidad de viviendas disminuye gradualmente. La distribución muestra asimetría y nos brinda información sobre la estructura y características del conjunto de datos de viviendas.

Se aprecia una distribución asimétrica en la que existen más viviendas en ciertos rangos de mayor superficie en comparación con otros rangos de menor superficie, aunque la diferencia con respecto al valor que más se repite sea la misma.

Es interesante notar que el histograma revela la escasez de viviendas con una superficie considerada grande. En estos rangos de superficie, la frecuencia o cantidad de viviendas es significativamente menor en comparación con los rangos de menor superficie. Esto indica una predominancia de viviendas con superficies más pequeñas y una disminución considerable en la cantidad de viviendas a medida que aumentan los metros cuadrados.

Figura 3: Distribución de los valores de las viviendas



Fuente: Elaboración propia

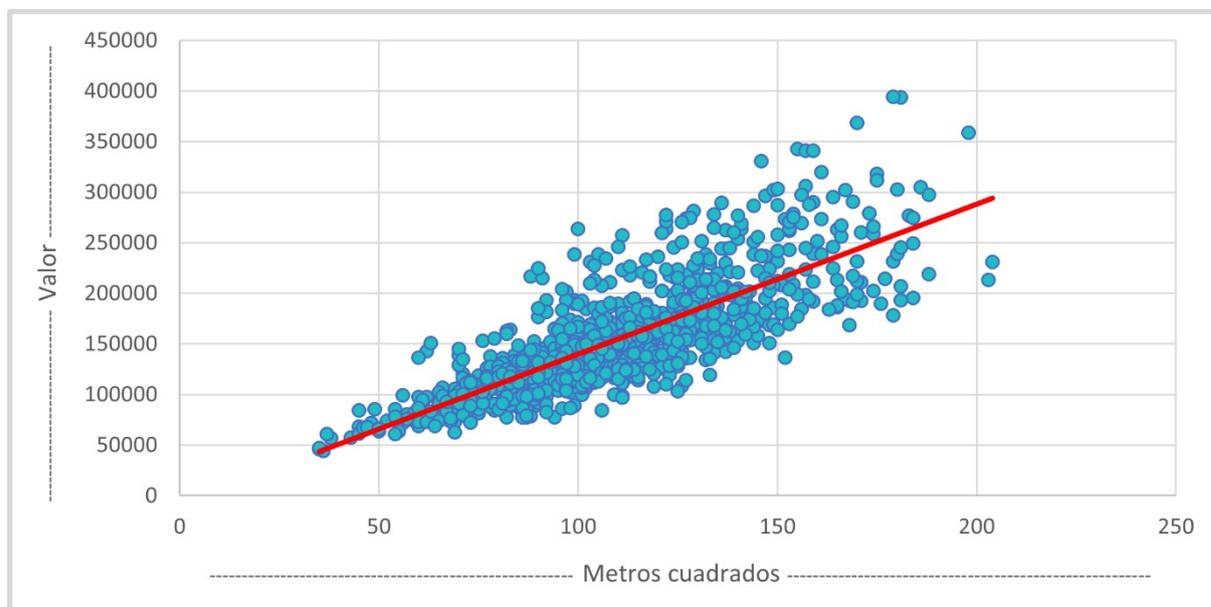
El histograma muestra la distribución del valor de las viviendas diferenciándolos por diferentes rangos. En el eje horizontal se representan estos rangos del valor de la vivienda, mientras que en el eje vertical se muestra la frecuencia o cantidad de viviendas que se encuentran dentro de cada rango.

Al observar el histograma, se puede apreciar que la mayoría de las viviendas se concentran en los rangos entre aproximadamente 120000 y 140000 euros. La frecuencia de viviendas es mayor en estos rangos, lo que indica que hay más viviendas con valores más económicos. Cuando baja de los 100000 euros se puede apreciar que la cantidad se reduce considerablemente comparado con los rangos de hasta 180000 euros que, aunque también se disminuye la cantidad, no se reduce de igual forma.

A medida que continuamos aumentando el valor de las viviendas, la frecuencia disminuye notablemente, lo que indica que hay menos viviendas con valores más altos. En el extremo derecho del histograma, se pueden observar algunos valores atípicos o viviendas con valores muy altos que se alejan significativamente de la media.

Estos valores atípicos representan viviendas con valores excepcionalmente altos en comparación con el resto de las viviendas, lo que sugiere la presencia de propiedades de lujo o con características especiales que las distinguen del resto del mercado inmobiliario.

Figura 4: Relación entre metros cuadrados y valor de vivienda: Regresión lineal



Fuente: Elaboración propia

En este gráfico se muestra una regresión lineal entre los metros cuadrados de las viviendas y su valor. Se puede observar que hay una tendencia positiva en la relación entre los metros cuadrados y el valor de las viviendas. A medida que aumenta la superficie de las viviendas, su valor tiende a ser más alto. Sin embargo, también se puede notar cierta dispersión alrededor de la línea de regresión, lo que indica que hay otros factores que influyen en el valor de las viviendas además de los metros cuadrados.

La regresión lineal nos permite estimar el valor de una vivienda en función de su superficie. Esta estimación se basa en la relación observada en los datos y nos proporciona una herramienta para predecir el valor de viviendas no incluidas en el conjunto de datos de entrenamiento.

En el gráfico de regresión lineal se observa que al final de la distribución, donde se encuentran las viviendas de mayor valor, se observa una mayor dispersión de los datos. Esto significa que existen otros factores que también influyen en el valor de estas viviendas y que no están siendo capturados únicamente por la superficie.

4 Desarrollo de un modelo de predicción de valores de bienes inmuebles utilizando redes neuronales

4.1 Análisis de datos generados

En el diseño y entrenamiento de la red neuronal se ha utilizado una arquitectura de red neuronal de tipo *MLPRegressor* (Regresor de Perceptrón Multicapa). Esta arquitectura se ha elegido por su capacidad para modelar relaciones no lineales entre las variables de entrada y la variable objetivo.

La elección de una red neuronal *feedforward*, como el *MLPRegressor*, puede ser idónea para el trabajo de predicción de valores de bienes inmuebles.

Las redes neuronales *feedforward* son capaces de capturar relaciones no lineales entre las variables de entrada y la variable objetivo. En el caso de los bienes inmuebles, es probable que existan relaciones complejas y no lineales

entre las características de una vivienda (como metros cuadrados, número de habitaciones, etc.) y su valor. El *MLPRegressor* permite modelar estas relaciones no lineales y capturar patrones complejos en los datos.

Las redes neuronales *feedforward* tienen una gran capacidad de aprendizaje y adaptabilidad. Pueden aprender a partir de datos complejos y descubrir relaciones sutiles entre las variables de entrada y la variable objetivo. En el caso de los bienes inmuebles, existen múltiples factores que pueden influir en el valor de una vivienda, como ubicación, características específicas, demanda del mercado, entre otros. Una red neuronal *feedforward* puede aprender a partir de estos datos complejos y capturar las relaciones relevantes para la predicción del valor de una vivienda.

Las redes neuronales *feedforward* pueden manejar conjuntos de datos con múltiples características y variables. En el caso de los bienes inmuebles, se pueden considerar varias características (metros cuadrados, número de habitaciones, baños, terraza, ascensor, parking, etc.) para predecir el valor de una vivienda. El *MLPRegressor* permite trabajar con múltiples características y aprender patrones complejos a partir de ellas.

El *MLPRegressor* permite ajustar la arquitectura de la red neuronal, como el número de capas ocultas, el número de neuronas por capa y la función de activación. Esto brinda flexibilidad para adaptar la red a la complejidad del problema y los datos disponibles. Es posible explorar diferentes configuraciones arquitectónicas para encontrar la que mejor se ajuste a los datos de bienes inmuebles.

La red neuronal está compuesta por dos capas ocultas, cada una con 100 neuronas. La función de activación utilizada en las capas ocultas es la función *ReLU* (*Rectified Linear Unit*), que permite introducir no linealidades en el modelo. El optimizador utilizado para el entrenamiento de la red es el Algoritmo Adam. Este algoritmo de optimización combina los conceptos del descenso de gradiente estocástico y el método de momentos, adaptando la tasa de aprendizaje de forma adaptativa mediante el uso de estimaciones de momentos de primer y segundo orden del gradiente.

Se optó por utilizar dos capas ocultas en la arquitectura de la red neuronal. Esta elección se basó en un equilibrio entre la capacidad de aprendizaje y la complejidad del modelo.

Utilizar más capas ocultas podría aumentar la capacidad de aprendizaje de la red neuronal, pero también podría llevar a un mayor riesgo de sobreajuste en conjuntos de datos más pequeños. Dos capas ocultas fueron consideradas suficientes para capturar relaciones no lineales complejas en los datos de bienes inmuebles.

Cada capa oculta se configuró con 100 neuronas. Esta elección se basó en experimentación y ajuste de la arquitectura. El número de neuronas por capa puede influir en la capacidad de aprendizaje y en la capacidad de generalización de la red neuronal. Demasiadas neuronas podrían llevar a un sobreajuste, mientras que muy pocas neuronas podrían limitar la capacidad de capturar patrones complejos en los datos. El número de 100 neuronas por capa fue seleccionado como un punto medio para abordar estas consideraciones.

Para preparar los datos de entrenamiento, se ha aplicado una estandarización utilizando el objeto *StandardScaler* de *scikit-learn*, con el fin de escalar las variables de entrada y centrarlas en torno a cero. La estandarización de los datos mediante el uso de *StandardScaler* asegura que las variables de entrada estén en una escala similar y centradas en torno a cero, lo cual es fundamental para el buen funcionamiento de la red neuronal y para obtener resultados más precisos y significativos en la predicción de los valores de los bienes inmuebles.

La arquitectura de la red neuronal se ha entrenado utilizando el conjunto de datos proporcionado en el archivo exportado de 10000 datos simulados. Se ha dividido el conjunto de datos en conjuntos de entrenamiento y prueba con un tamaño de prueba del 20% del tamaño total de los datos, mientras que el conjunto de entrenamiento representa el 80% restante. Al dividir el conjunto de datos en conjuntos de entrenamiento y prueba, es importante mantener la reproducibilidad de los resultados. Para lograr esto, se utiliza una semilla aleatoria que garantiza que la división en conjuntos sea la misma cada vez que se ejecute el código, en este caso se ha utilizado el número 42 como semilla aleatoria.

Una vez entrenada la red neuronal, se realiza una predicción sobre una nueva vivienda. Estos valores de la nueva vivienda se ajustan a los tipos de datos esperados y se escalan utilizando el mismo objeto *StandardScaler* utilizado durante el entrenamiento.

4.2 Evaluación del modelo

La evaluación del modelo de red neuronal se basa en el análisis de varias métricas y resultados obtenidos durante el proceso. Esta evaluación se realizó utilizando un conjunto de prueba que representa el 20% del tamaño total de los datos, es decir, 2000 datos, mientras que el conjunto de entrenamiento comprende el 80% restante.

Al evaluar la red neuronal, se utilizaron varias métricas para comprender su rendimiento y precisión en la predicción de los valores de los bienes inmuebles.

El coeficiente de determinación (R^2) se obtuvo con un valor de 0,9771 en el conjunto de prueba. Este coeficiente proporciona una medida de la bondad de ajuste del modelo a los datos. Un valor más cercano a 1 indica un mejor ajuste del modelo, y en este caso, el valor alto del coeficiente R^2 sugiere que aproximadamente el 97.71% de la variabilidad de los valores de los bienes inmuebles en el conjunto de prueba puede ser explicada por el modelo.

El error absoluto medio se calculó con un valor de 5880,77 unidades monetarias. Esta métrica mide el promedio de los errores absolutos entre los valores reales y las predicciones del modelo. El error absoluto medio proporciona una medida de lo cerca que están las predicciones del modelo de los valores reales, sin considerar la dirección del error. En este caso, el error absoluto medio indica que, en promedio, las predicciones del modelo tienen un error absoluto de aproximadamente 5880.77 unidades monetarias.

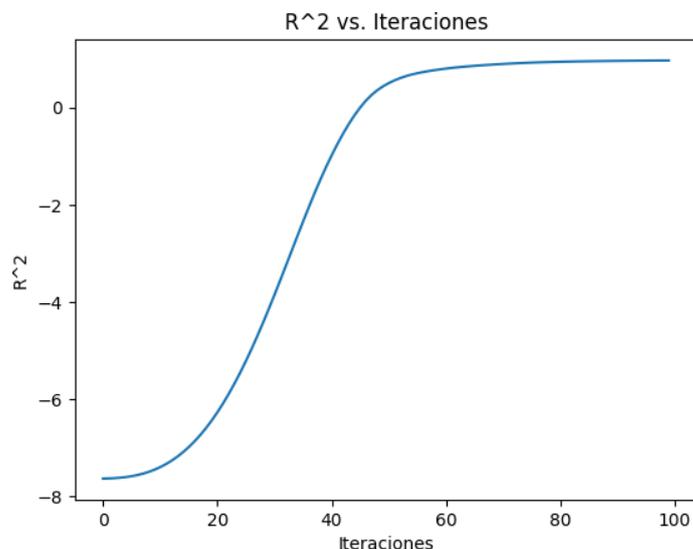
El error cuadrático medio se calculó con un valor de 70249078,39 unidades monetarias al cuadrado. Esta métrica mide el promedio de los errores al cuadrado entre los valores reales y las predicciones del modelo. El error cuadrático medio amplifica los errores más grandes y puede ser útil para evaluar la dispersión de los errores. En este caso, el valor del error cuadrático medio indica la magnitud de los errores cuadráticos promedio.

La raíz del error cuadrático medio se obtuvo con un valor de 8381,47 unidades monetarias. Esta métrica se calcula como la raíz cuadrada del error cuadrático medio y proporciona una medida de dispersión que indica qué tan cerca están las predicciones del modelo de los valores reales. En este caso, la raíz del error cuadrático medio indica que, en promedio, las predicciones del modelo se desvían aproximadamente en 8381,47 unidades monetarias de los valores reales.

En conclusión, la red neuronal ha demostrado un buen rendimiento en la predicción de los valores de los bienes inmuebles. El alto valor del coeficiente de determinación (R^2) indica un buen ajuste del modelo a los datos, y las métricas de error proporcionan información sobre la precisión y dispersión de las predicciones. Estas métricas y resultados nos permiten tener una comprensión completa del rendimiento del modelo en la tarea de predicción de valores de bienes inmuebles.

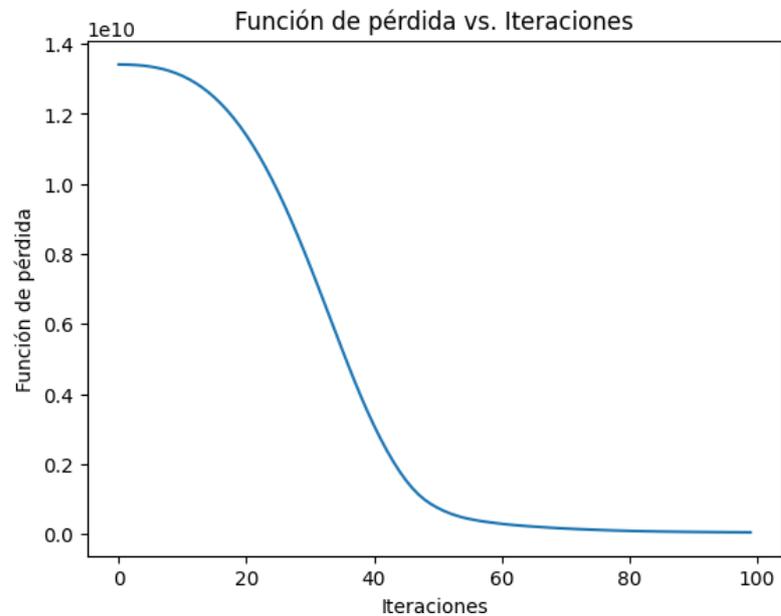
A continuación, se presentan los gráficos generados durante la evaluación del modelo:

Figura 5: Gráfico del coeficiente de determinación



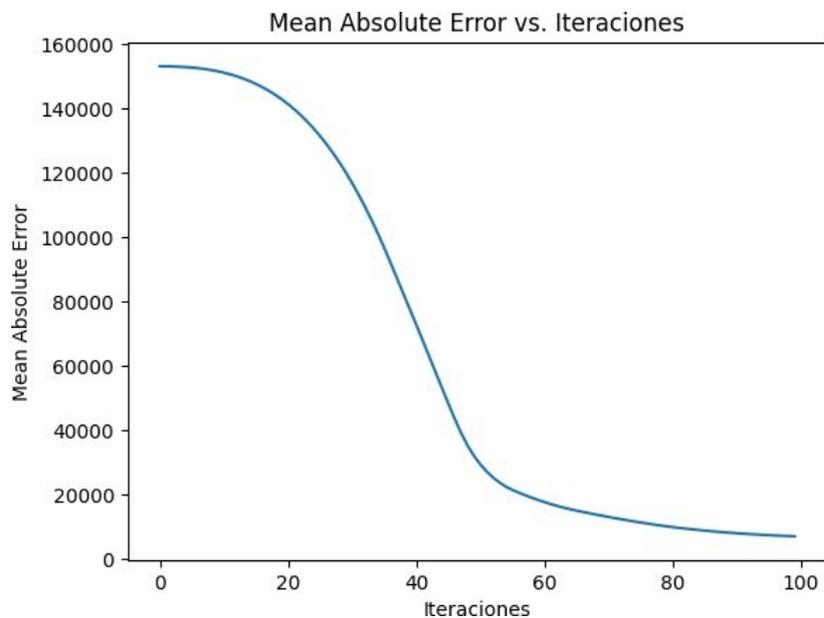
En este gráfico, se observa cómo el valor del coeficiente de determinación se acerca gradualmente a 1 a medida que se realizan más iteraciones de entrenamiento. Esto indica una mejor capacidad del modelo para explicar la variabilidad de los datos.

Figura 6: Gráfico de la función de pérdida



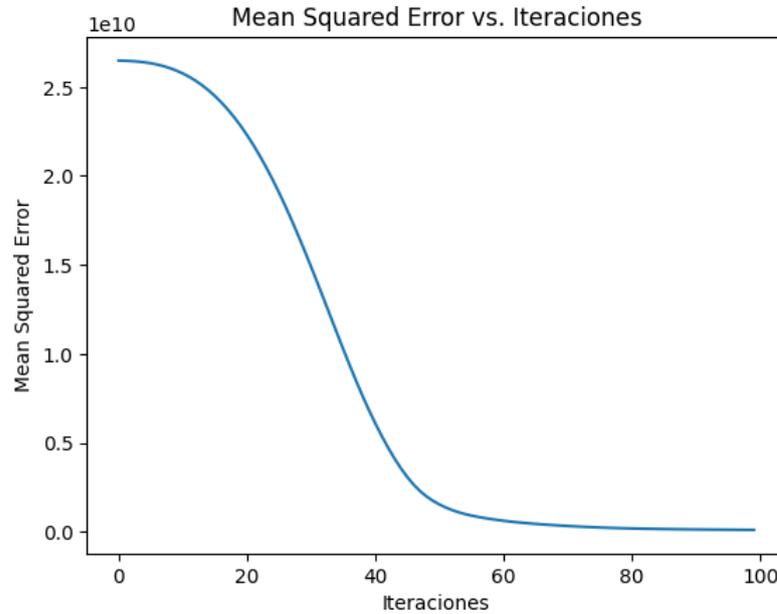
Este gráfico muestra cómo disminuye la función de pérdida a medida que aumentan las iteraciones de entrenamiento. La disminución de la función de pérdida indica una mejora continua del modelo.

Figura 7: Gráfico del error absoluto medio



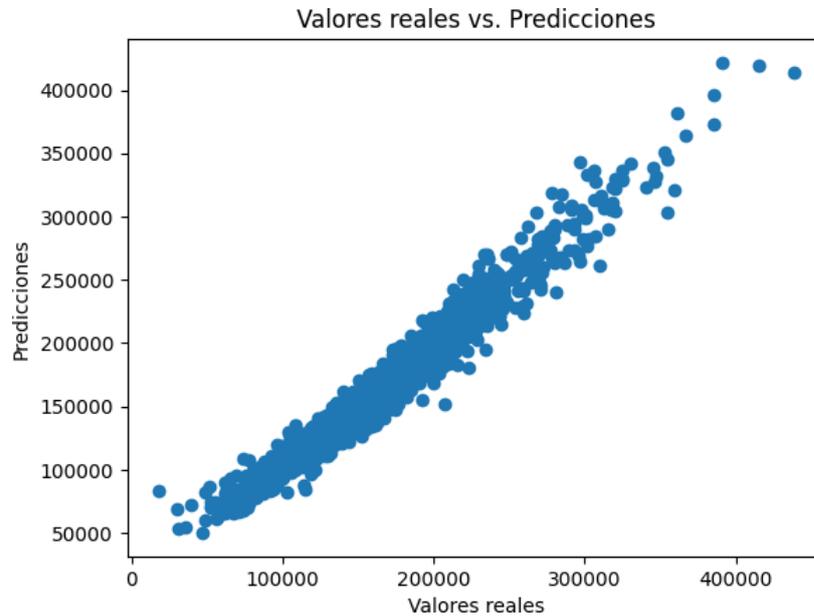
Este gráfico muestra cómo el error absoluto medio disminuye a medida que se realizan más iteraciones de entrenamiento, lo cual sugiere una mejora en la precisión del modelo.

Figura 8: Gráfico del error cuadrático medio



En este gráfico, se puede apreciar cómo el error cuadrático medio disminuye a medida que se realizan más iteraciones de entrenamiento, lo que indica una reducción en la dispersión de los errores.

Figura 9: Dispersión de valores reales vs predicciones



Este gráfico permite visualizar la comparación entre las predicciones del modelo y los valores reales en el conjunto de prueba. Idealmente, las predicciones deberían estar cerca de la línea diagonal, lo que indicaría una correspondencia cercana con los valores reales. Esto es lo que sucede en el modelo planteado.

4.3 Ejemplos de cálculo del valor de viviendas

Tabla 3: Ejemplos valores predichos por la red neuronal

Barrio	C. Jardín - Zoco	Fátima - Levante	Zona Centro	Sector Sur
M²	100	100	110	70
Nº Habitaciones	4	4	4	3
Nº Baños	2	2	2	1
Terraza	Sí	Sí	Sí	Sí
Ascensor	Sí	Sí	Sí	Sí
Parking	No	No	Sí	No
Amueblado	No	No	Sí	Sí
Valor (€)	156295	132934	225674	78830

Fuente: Elaboración propia

La tabla presentada muestra cuatro ejemplos del poder predictivo de una red neuronal entrenada con un conjunto de datos de ventas de viviendas.

Una vez completado el entrenamiento con esos datos, la red neuronal está lista para realizar predicciones precisas del valor de venta de una vivienda en función de los datos específicos proporcionados. En estos ejemplos, hemos seleccionado cuidadosamente diferentes conjuntos de características deseadas para cada vivienda, como la ubicación en un barrio específico y los metros cuadrados deseados.

Estos ejemplos proporcionan una visión esclarecedora de cómo la red neuronal puede aprovechar su conocimiento adquirido durante el entrenamiento para realizar predicciones perspicaces y útiles para el mercado inmobiliario.

Las dos primeras columnas muestran cómo, para los mismos datos de entrada, el valor de venta de las viviendas varía según el barrio. Esto demuestra que otros aspectos, como la ubicación y las características específicas de cada barrio, influyen de manera notable en el precio de la vivienda.

En las dos siguientes columnas, se destaca la diferencia en el valor de venta entre dos viviendas. Una de ellas se encuentra en la Zona Centro, un área reconocida por tener precios elevados en comparación con la mayoría de las zonas de la ciudad de Córdoba. Por otro lado, la otra vivienda se sitúa en el Sector Sur, considerado como una de las zonas más asequibles en términos de precios de vivienda en la ciudad.

Estos ejemplos ilustran cómo la red neuronal utiliza la información proporcionada por los datos de entrada para estimar el valor de venta de las viviendas en distintos barrios.

Su capacidad para captar patrones complejos y sutilezas en los datos permite generar estimaciones confiables que reflejan la influencia de diversos factores en el valor de una vivienda. Esto proporciona una valiosa herramienta para tomar decisiones informadas en el mercado inmobiliario y maximizar la precisión en las estimaciones de precios de venta de viviendas.

5 Conclusiones

En este Trabajo de Fin de Grado, se abordó el tema de las aplicaciones de las redes neuronales en el mercado inmobiliario en Córdoba. Uno de los desafíos significativos fue la limitada disponibilidad de datos para el entrenamiento de la red neuronal. A diferencia de ciudades más grandes y reconocidas como Madrid o Barcelona, la información detallada y completa que hemos intentado buscar y acceder sobre el mercado inmobiliario en Córdoba es escasa.

La disponibilidad de información precisa y completa en el mercado inmobiliario puede ser un desafío en muchas situaciones. En algunos casos, los datos relevantes pueden estar protegidos por la Ley de Protección de Datos, ser confidenciales de las inmobiliarias o requerir un pago para su acceso.

Esta falta de datos representativos ha supuesto un obstáculo importante en el desarrollo del estudio, ya que los modelos de redes neuronales requieren conjuntos de datos amplios y actualizados para obtener resultados óptimos. Sin embargo, se ha abordado este problema mediante la simulación de datos a partir de datos en torno a medias como el precio medio por metro cuadrado diferenciado por barrios o la superficie que más se repite en las ventas de viviendas de esos barrios concretos. Aunque estos datos simulados no son observaciones reales, han permitido avanzar en el análisis y entrenamiento de la red neuronal.

A pesar de las limitaciones asociadas a la simulación de datos, los resultados obtenidos a través del entrenamiento de la red neuronal han mostrado un desempeño positivo y cercano a la realidad. Las estimaciones generadas se han aproximado de manera significativa a los valores reales del mercado inmobiliario en Córdoba. Esto indica que, a pesar de la falta de datos concretos, las redes neuronales pueden ser una herramienta eficaz para la predicción y valoración de propiedades en Córdoba.

No obstante, es importante reconocer que la falta de datos reales representa una limitación en este estudio. La ausencia de información precisa puede generar cierta incertidumbre en los resultados y afectar su validez en comparación con investigaciones a partir de datos reales. Además, la extrapolación de estos resultados a otras ciudades o contextos del mercado inmobiliario debe realizarse con precaución, ya que las características específicas de cada mercado pueden influir en la aplicabilidad de los modelos.

Una de las contribuciones significativas de este estudio es la creación de un código escrito en Python capaz de llevar a cabo una simulación de datos en ausencia de información real para el entrenamiento de la red neuronal. Esta aproximación ha permitido superar la limitación inicial de la falta de datos disponibles y ha proporcionado una base sólida para el desarrollo del modelo.

A través del proceso de entrenamiento de la red neuronal con los datos simulados, se han obtenido estimaciones que se acercan de manera significativa a los valores reales en el mercado inmobiliario cordobés. Este resultado demuestra el potencial de las redes neuronales para ofrecer predicciones precisas en este contexto y respalda su utilidad como herramienta de valoración de propiedades. Esta red neuronal también se ha elaborado de forma personalizada capaz de trabajar con los datos simulados previamente.

La generación de datos simulados a partir de un código original de Python es una contribución original de este trabajo. Este enfoque puede ser de interés para investigadores y profesionales del sector, ya que abre nuevas posibilidades en la construcción de conjuntos de datos representativos cuando la información real es escasa o no está disponible.

Las conclusiones y metodologías desarrolladas en este trabajo pueden tener aplicabilidad más amplia ya que, aunque este estudio se centra en el mercado inmobiliario en Córdoba, sus hallazgos y enfoques pueden ser extrapolados a otros contextos y ciudades.

Este trabajo no solo contribuye al campo de las redes neuronales aplicadas al mercado inmobiliario, sino que también proporciona una solución práctica para el desafío específico de la escasez de información en el contexto de Córdoba o en ciudades con características similares, donde la información disponible es limitada y se requieren enfoques alternativos para el análisis y la toma de decisiones.

Referencias

- Abeijón, D. (2007). *Técnicas de fusión de datos*. Descargado de <https://upcommons.upc.edu/bitstream/handle/2099.1/4376/03.pdf?sequence=28&isAllowed=y>
- Barrios Arce, J. I. (2023, 5 de marzo). Generación de datos sintéticos: una solución eficiente para el problema de datos insuficientes. Descargado de <https://www.juanbarrios.com/generacion-de-datos-sinteticos-el-problema-de-datos-insuficientes/>

- Cambios en la compra de vivienda por la covid-19.* (2021, 24 de febrero). Estardondeestes. Descargado de <https://estardondeestes.com/movi/es/articulos/cambios-en-la-compra-de-vivienda-por-la-covid-19>
- Canal Díaz, N. (s.f.). Técnicas de muestreo. sesgos más frecuentes. Descargado de <https://revistaseden.org/files/9-CAP%209.pdf>
- Chaipimonplina, T., y Sin-Ampola, P. (2019). *Capacity building for research and innovation in disaster resilience.* Descargado de <https://sparc.cmb.ac.lk/wp-content/uploads/2019/07/Final-version-of-the-ASCENT-Book-volume-.pdf>
- Choy, L. H. T., y Ho, W. K. O. (2023). The use of machine learning in real estate research. *Land*. Descargado de <https://www.mdpi.com/2073-445X/12/4/740>
- C.L. (2023, 9 de febrero). ¿cuánto vale una plaza de garaje en córdoba? *El Día de Córdoba*. Descargado de https://www.eldiadecordoba.es/provincia/cuanto-vale-plaza-garaje-Cordoba_0_1764425588.html
- DataScientest. (2022a, 25 de enero). *Random forest: Bosque aleatorio. definición y funcionamiento.* Descargado de <https://datascientest.com/es/random-forest-bosque-aleatorio-definicion-y-funcionamiento>
- DataScientest. (2022b, 6 de enero). *¿qué es el transfer learning?* Descargado de <https://datascientest.com/es/que-es-el-transfer-learning>
- Decision trees. (s.f.). Descargado de <https://www.ibm.com/es-es/topics/decision-trees>
- Diferencia entre datos estructurados y no estructurados. (s.f.). Descargado de <https://www.kyoceradocumentsolutions.es/es/smarter-workspaces/insights-hub/articles/diferencia-entre-datos-estructurados-y-no-estructurados.html>
- Di Pietro, M. (2021, 17 de diciembre). Deep learning with python neural networks: Complete tutorial. Descargado 2023-03-07, de <https://towardsdatascience.com/deep-learning-with-python-neural-networks-complete-tutorial-6b53c0b06af0>
- Durán, J. (2020, 22 de enero). Qué es la transferencia de aprendizaje y cómo aplicarla a tu red neuronal. Medium. Descargado de <https://medium.com/metadatos/qu%C3%A9-es-la-transferencia-de-aprendizaje-y-c%C3%B3mo-aplicarla-a-tu-red-neuronal-e0e120156e40>
- Ermakovich, S. (2023, 27 de febrero). Web scraping vs api: What's the best way to extract data? Descargado de <https://scrape-it.cloud/blog/web-scraping-vs-api>
- García Navia, A. (2022, 26 de abril). Mercado inmobiliario: qué elementos suben el precio de una vivienda. TheObjective. Descargado de <https://theobjective.com/economia/2022-04-26/vivienda-precio-elementos/>
- Gill, N. S. (2023, 15 de enero). Artificial neural networks applications and algorithms. Descargado de <https://www.xenonstack.com/blog/artificial-neural-network-applications>
- Great Learning Team. (2022, 23 de noviembre). Types of neural networks and definition of neural network. Descargado de <https://www.mygreatlearning.com/blog/types-of-neural-networks/>
- Introducción a las redes neuronales recurrentes. (2019, 8 de junio). Descargado de <https://www.codificandobits.com/blog/introduccion-redes-neuronales-recurrentes/#algunas-aplicaciones-de-las-redes-neuronales-recurrentes>
- Kalliola, J., Kapočiūtė-Dzikienė, J., y Damaševičius, R. (2021). Neural network hyperparameter optimization for prediction of real estate prices in helsinki. PeerJ. Computer science.
- Kaushik, V. (2021, 27 de agosto). *8 applications of neural networks.* Descargado de <https://www.analyticssteps.com/blogs/8-applications-neural-networks>
- KeepCoding, R. (2022, 19 de enero). *Las 8 herramientas de data analytics más usadas.* Descargado de <https://keepcoding.io/big-data/herramientas-de-data-analytics-mas-usadas/>

- Khalafallah, A. (2008, octubre). Neural network based model for predicting housing market performance. Tsinghua Science Technology. Descargado de https://www.researchgate.net/publication/245476723_Neural_Network_Based_Model_for_Predicting_Housing_Market_Performance
- Malingan, N. (2023, 4 de Mayo). Transfer learning and fine tuning. Descargado de <https://www.scaler.com/topics/deep-learning/transfer-learning-and-fine-tuning/>
- Mesquita, D. (2021). Python ai: How to build a neural network & make predictions. Descargado de <https://realpython.com/python-ai-neural-network/>
- Mouzo, F. (2022, 3 de abril). ¿cuántas viviendas en venta o alquiler en córdoba tienen ascensor? El Día de Córdoba. Descargado de https://www.eldiadicordoba.es/cordoba/viviendas-venta-alquiler-Cordoba-ascensor_0_1670234173.html
- Neural networks. (s.f.). Descargado de <https://www.ibm.com/topics/neural-networks>
- OneClick, G. (2021, 20 de octubre). Limpieza y Normalización: los 2 métodos clave para mejorar la Calidad de tus Datos. Descargado de <https://www.linkedin.com/pulse/limpieza-y-normalizaci%C3%B3n-los-2-m%C3%A9todos-clave-para-mejorar-/?originalSubdomain=es>
- Paredes, D. (2020, 10 de marzo). Introducción a autoencoders. Descargado de <https://atcold.github.io/pytorch-Deep-Learning/es/week07/07-3/>
- Perez, M. (2023, 14 de abril). What is web scraping and what is it used for? Descargado de <https://www.parsehub.com/blog/what-is-web-scraping/>
- Precio viviendas córdoba capital m2 - mayo de 2023. (2023). Descargado de <https://www.fotocasa.es/indice-precio-vivienda/cordoba-capital/todas-las-zonas>
- Rayón, A. (2015, 23 de octubre). Eligiendo una herramienta de analítica: Sas, r o python. Descargado de <https://blogs.deusto.es/bigdata/eligiendo-una-herramienta-de-analitica-sas-r-o-python/>
- RB, A., y KR, S. K. (2021). Credit card fraud detection using artificial neural network. Global Transitions Proceedings. Descargado de <https://www.sciencedirect.com/science/article/pii/S2666285X21000066>
- Ruiz, A. S. (2022, 23 de agosto). España ocupa el puesto 18 en transparencia en el mercado inmobiliario. Cinco Días. Descargado de https://cincodias.elpais.com/cincodias/2022/08/22/companias/1661190181_392476.html
- Saha, S. (2018, 15 de diciembre). A comprehensive guide to convolutional neural networks — the eli5 way. Descargado de <https://saturncloud.io/blog/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/>
- Sólo el 35% de las viviendas en venta en España tiene plaza de garaje. (2017, 14 de junio). Idealista. Descargado de <https://www.idealista.com/news/inmobiliario/vivienda/2017/06/14/746871-solo-el-35-de-las-viviendas-en-venta-en-espana-tienen-plaza-de-garaje#:~:text=Pr%C3%A1cticamente%20%20de%20cada%203.el%20marketplace%20inmobiliario%20de%20Espa%C3%BAa.>
- Tiet, K. H. (2020, 28 de septiembre). When does transfer learning fail in deep learning. Descargado de <https://medium.com/aviation-software-innovation/when-does-transfer-learning-fail-in-deep-learning31bc2c93f18a>
- Turing. (2022). Understanding feed forward neural networks with maths and statistics. Online. Descargado de <https://www.turing.com/kb/mathematical-formulation-of-feed-forward-neural-network>
- UNIR. (2022, 18 de abril). Tipos de muestreo: los principales y sus características. Revista de Ingeniería. Descargado de <https://www.unir.net/ingenieria/revista/tipos-de-muestreo/>

- Universitat de València. (s.f.). Cluster. Descargado de <https://www.uv.es/ceaces/multivari/cluster/CLUSTER2.htm>
- Vaquero, J. (2022, 17 de abril). La pandemia ha disparado al sector inmobiliario que ha empezado a evolucionar muy bien. Cordoba Hoy. Descargado de <https://www.cordobahoy.es/articulo/gente/entrevista/20220415213414114684.html>
- Viana Colino, M. (2012). Minería de datos en portal inmobiliario (Trabajo de Fin de Grado). Descargado de https://e-archivo.uc3m.es/bitstream/handle/10016/16683/PFC_Marta_Viana_Colino.pdf;jsessionid=F1A775FEBE3ED01362512CCF709CBCB7?sequence=1
- Walczak, S. (2021). Frontiers in Psychology. Descargado de <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.587943>
- What is a random forest? (s.f.). Descargado de <https://www.tibco.com/es/reference-center/what-is-a-random-forest>
- ¿cuánto ha subido el precio de la vivienda en córdoba? (2023, 22 de marzo). Descargado de https://www.eldiadicordoba.es/cordoba/subido-precio-vivienda-Cordoba_0_1777322594.html
- ¿qué es la regresión lineal? (s.f.). Amazon Web Services, Inc. Descargado de <https://aws.amazon.com/es/what-is/linear-regression/>