

Universidad de Buenos Aires
Facultad de Ciencias Económicas
Escuela de Negocios y Administración Pública

**CARRERA DE ESPECIALIZACIÓN EN
MÉTODOS CUANTITATIVOS PARA LA GESTIÓN Y ANÁLISIS
DE DATOS EN ORGANIZACIONES**

TRABAJO FINAL INTEGRADOR

Análisis del Impacto de la Pandemia en el
Mercado mediante Modelos de Aprendizaje
Automático

AUTOR: DOMINIQUE MAINHARD
TUTOR: ROBERTO ABALDE

SEPTIEMBRE 2024

Resumen

La pandemia de coronavirus transformó la vida de personas en todo el mundo. Durante más de un año, los ciudadanos argentinos se vieron obligados a continuar con sus vidas y rutinas dentro de sus hogares, sin importar la situación en la que se encontraran. Actividades como trabajar, estudiar, hacer ejercicio e incluso disfrutar del ocio debieron realizarse dentro de casa. Este nuevo panorama impactó directamente en la perspectiva de las personas, llevándolas a revisar sus prioridades y necesidades. El presente trabajo se centra en las consecuencias específicas que esta situación tuvo en el mercado inmobiliario.

Buenos Aires Real Estate es una plataforma de clasificados inmobiliarios que será la principal beneficiaria de los resultados de este estudio. El objetivo es mejorar la precisión de las herramientas que ofrece para la predicción de precios de inmuebles, empleando técnicas avanzadas de ciencias de datos. Este trabajo busca, además, identificar y analizar las variables con mayor influencia en el valor del metro cuadrado antes y después de la pandemia, así como detectar las tendencias que podrían dominar en el futuro.

El objetivo general es identificar las variables con mayor impacto en el precio de los inmuebles en Buenos Aires en la actualidad, comprender su incidencia y, posteriormente, predecir los precios en función de estas variables.

Hasta hace unos años, se podía afirmar que la ubicación y la cantidad de metros cuadrados cubiertos eran las variables con mayor incidencia en el precio de cualquier inmueble, en cualquier parte del mundo. Este trabajo analiza cómo, tras la pandemia, variables que antes no tenían un papel primordial, como la de metros descubiertos, han ganado especial relevancia.

La hipótesis principal sostiene entonces, que ciertas variables -anteriormente secundarias- han cobrado mayor protagonismo y hoy en día influyen más en el precio de los inmuebles que hace tres años. El enfoque será cuantitativo, recopilando datos históricos de propiedades en venta en Capital Federal entre 2020 y 2024. Se emplearán técnicas de análisis estadístico y modelos de aprendizaje automático, comparando algoritmos avanzados con métodos tradicionales como la

regresión lineal. La precisión y aplicabilidad del modelo predictivo se evaluarán utilizando métricas como el error cuadrático medio (RMSE).

Este trabajo tiene como objetivo proporcionar una herramienta valiosa para el mercado inmobiliario local, mejorando la transparencia y accesibilidad de la información para todos los ciudadanos de Buenos Aires.

Palabras clave: Mercado inmobiliario, Predicción de precios, Aprendizaje automático, Buenos Aires, Pandemia

Indice

Resumen	2
Introducción	5
1. El Mercado Inmobiliario en la actualidad	7
1.1. Pre y post pandemia	7
1.2. La ubicación como variable principal	11
1.3. La integración tecnológica en el mercado inmobiliario	13
2. Construcción y Análisis de la Base de Datos	14
2.1. Selección de Datos	14
2.2. Limpieza de Datos	15
2.3. Análisis Exploratorio	16
3. El Rol de los Modelos Predictivos en las Plataformas Digitales	22
3.1. Desarrollo de modelos de predicción	23
3.2. Optimización y Resultados	24
3.3. Aplicación de modelos de predicción	27
Conclusiones	28
Referencias bibliográficas	30
Apéndices	33
Apéndice I - Conjunto de Datos utilizado para el período pre-pandemia 2020-2021:	33
Apéndice II - Conjunto de Datos utilizado para el período post-pandemia 2023-2024:	33
Apéndice III - Cuaderno Colab con la limpieza de datos y armado de modelo del período pre-pandemia (2020-2021):	33
Apéndice IV - Cuaderno Colab con la limpieza de datos y armado de modelo del período post-pandemia (2023-2024):	33
Reporte del Mentor	34

Introducción

La pandemia de coronavirus se convirtió en una de las crisis globales más significativas de la historia reciente, y el sector inmobiliario fue notablemente afectado. La industria inmobiliaria enfrentó grandes desafíos debido a la incertidumbre económica y las restricciones de distanciamiento social.

Este trabajo tiene como objetivo analizar el impacto de la pandemia en el mercado inmobiliario de Buenos Aires hacia el año 2024.

Uno de los cambios más notables derivados de la pandemia fue la adopción masiva del trabajo remoto, o *home office*. Antes de la crisis sanitaria, el trabajo desde casa era una opción para pocos, pero las restricciones de movilidad y las medidas de confinamiento forzaron a muchas empresas y empleados a adaptarse a esta modalidad. Esta transformación abrupta en la forma de trabajar reveló nuevas necesidades en el entorno doméstico. La necesidad de espacios de trabajo funcionales dentro del hogar se convirtió en una prioridad, impulsando a los individuos a buscar propiedades que ofrecieran áreas adecuadas para el teletrabajo.

El trabajo remoto no solo modificó la forma en que se organizan las oficinas, sino que también cambió la percepción de lo que constituye un hogar ideal. La importancia de contar con un ambiente separado para el trabajo, libre de distracciones y adecuadamente equipado, se hizo evidente. Además, el confinamiento prolongado expuso la necesidad de espacios que promuevan el bienestar mental y físico, haciendo que los espacios verdes y las áreas exteriores se convirtieran en elementos esenciales en la elección de un lugar para vivir.

La presión de adaptar los hogares para el trabajo remoto llevó a una reevaluación de la funcionalidad y el diseño de los espacios residenciales. Las propiedades con áreas al aire libre, como jardines o terrazas, y las que se encontraban cerca de parques o espacios verdes, empezaron a ser más valoradas. La proximidad a estos espacios se convirtió en un factor importante, ya que ofrecían oportunidades para actividades recreativas y de esparcimiento, esenciales para contrarrestar el estrés asociado con el trabajo y el confinamiento.

Como resultado de estas nuevas prioridades, el mercado inmobiliario experimentó un cambio en la demanda. Variables que antes eran secundarias, como la disponibilidad de metros cuadrados descubiertos y la cercanía a espacios verdes, comenzaron a tener un impacto significativo en las decisiones de compra y en la valoración de las propiedades. Este cambio en las necesidades residenciales refleja una tendencia más amplia hacia la búsqueda de viviendas que ofrezcan un equilibrio entre el espacio para el trabajo y áreas que contribuyan al bienestar general.

Actualmente, no existe una claridad absoluta sobre las variables más influyentes en el precio de los inmuebles en Buenos Aires, lo que complica la predicción precisa de precios. Este trabajo busca demostrar cómo factores como los metros descubiertos y la cercanía a zonas verdes han ganado protagonismo en el mercado inmobiliario post pandemia.

Se desarrollará un modelo predictivo para evaluar si estas características han aumentado su peso en el mercado, superando a variables tradicionales como la cantidad de baños.

Las variables mencionadas han ganado relevancia debido a una mayor conciencia global sobre posibles futuras pandemias. La importancia de características como los espacios verdes marca un cambio en el mercado inmobiliario. Este estudio pretende analizar cómo estas variables han influido en la evolución de los precios de los inmuebles antes y después de la pandemia.

La relevancia de esta investigación radica en su capacidad para proporcionar información valiosa sobre las variables que afectan los precios de venta de los inmuebles en diferentes contextos. *Buenos Aires Real Estate*, una plataforma en línea de clasificados inmobiliarios, será la principal beneficiaria, al mejorar la transparencia y accesibilidad del mercado a través de herramientas más precisas y eficaces.

Este estudio explora la posibilidad de desarrollar un modelo predictivo preciso utilizando métodos cuantitativos y datos históricos. Se evaluará el impacto de variables como la ubicación y las características de los inmuebles en el precio de venta. La incorporación de técnicas avanzadas de aprendizaje automático representa una innovación en el sector inmobiliario, contribuyendo al avance del mercado y mejorando la transparencia.

Las preguntas centrales de este trabajo son: ¿Cómo han cambiado los factores determinantes del valor de las propiedades en Buenos Aires antes y después de la pandemia? ¿Cuál es la relevancia actual de los espacios descubiertos en la determinación de precios? ¿Es posible predecir con precisión los precios utilizando métodos cuantitativos? Estos interrogantes guían los objetivos principales de la investigación:

- * Identificar las variables con mayor incidencia en el precio de los inmuebles en Buenos Aires, y comprender la incidencia de la pandemia por coronavirus en la ponderación de las mismas.
- * Explorar la correlación entre los metros descubiertos y el precio de los inmuebles en Buenos Aires.
- * Predecir la variable precio según las variables identificadas.

Las hipótesis propuestas incluyen:

1. Variables como amenities, gimnasio, parrilla, pileta o espacio de cowork han revalorizado los precios de los inmuebles y hoy tienen mayor incidencia que antes.
2. La cercanía a zonas verdes incide más en el precio que el acceso a ubicaciones céntricas.
3. Después de la pandemia, el precio por metro cuadrado descubierto ha ganado relevancia.

El análisis se estructura en tres capítulos: el primero describe el panorama del mercado inmobiliario argentino y los efectos de la pandemia; el segundo se enfoca en las variables más relevantes y su influencia en el precio final; y el tercero examina el impacto de la tecnología en el sector inmobiliario, con especial atención a *Buenos Aires Real Estate*. Finalmente, se presentan las conclusiones, respondiendo a los interrogantes planteados al inicio del estudio.

1. El Mercado Inmobiliario en la actualidad

El primer apartado explora las tendencias emergentes en la industria inmobiliaria tras la pandemia, que marcó un antes y un después en el sector. La vida de las personas se transformó

con la adopción de nuevos hábitos y necesidades, cambios que se reflejan directamente en el mercado inmobiliario.

1.1. Pre y post pandemia

La principal medida para mitigar la pandemia fue el aislamiento o cuarentena de la población, denominado Aislamiento Social Preventivo y Obligatorio (ASPO). Dicha medida fue anunciada el 20 de marzo del 2020 y tuvo vigencia hasta el 9 de noviembre del mismo año. Durante 234 días, todos los habitantes de la Argentina se vieron obligados a quedarse puertas adentro.

Durante casi un año, las personas debieron adaptarse para continuar con sus vidas dentro de sus casas. Claro que no fue tarea fácil, y a algunos esta situación los encontró en escenarios más o menos favorables. De un día para el otro, las personas debieron encontrar la forma de poder trabajar dentro de sus casas, los jóvenes estudiar en simultáneo, mientras que encontraban la forma de pasar sus momentos de ocio, de hacer actividad física y mantener los vínculos sociales. Todo esto sucedía puertas adentro.

Luego de la pandemia la casa se convirtió más que nunca en un refugio. Un lugar seguro. En este sentido, el mercado inmobiliario comenzó a notar los cambios en las tendencias de demanda. La pandemia generó un impacto enorme en las nuevas formas de vivir, profundizando así nuevas necesidades y deseos en torno al lugar que los argentinos elegían para vivir. De hecho, 7 de cada 10 afirman que a partir de la crisis sanitaria evaluaron la necesidad de mudarse, según reveló una encuesta de Mercado Libre Inmuebles (2020). Además, entre los que decían estar "insatisfechos" con su vivienda ante la pandemia, el motivo más mencionado -dicho por un 62%- era la falta de espacio, ya sea interior (38%) o exterior (24%).

Una década atrás, los departamentos con más metros cuadrados cubiertos eran preferidos, especialmente aquellos con amplias habitaciones y salones. Según una nota del diario Clarín (2021), en el último tiempo la idea fija pasó a ser sumar metros cuadrados descubiertos. El informe indica que “la gran obsesión es que el inmueble tenga algún acceso al aire libre, llámese

balcón, terraza, patio o jardín. Aun si para eso deben renunciar a otras características antes más valoradas.”

Los departamentos con buenas vistas y diseños modernos o renovados también gozaban de alta apreciación. Dado el contexto cambiante y los nuevos requisitos emergentes, es posible que estas prioridades y preferencias en el mercado inmobiliario hayan cambiado. En este contexto previo a la pandemia, la tendencia pasaba por comprar o alquilar en edificios con amenities, como gimnasio, sauna, piscina y SUM. Sin embargo, con la extensión de las medidas restrictivas y el distanciamiento social hoy la demanda está más enfocada en elegir viviendas con balcones amplios, patios y terrazas.

En el mismo sentido, la demanda de casas y terrenos con espacios verdes comenzó a ganar espacio, confirmando una tendencia que comenzó a mediados de 2020 y que sigue creciendo, afirma Mudafy (2022). Tal es así que durante buena parte del 2021, según Daniel Bryn, especialista en el mercado inmobiliario, por cada 100 personas interesadas en comprar en 2021, 63 priorizaron casas, 21 se focalizaron en terrenos y 10 optaron por propiedades horizontales o PH. En palabras de Bryn, “el consumidor hoy prioriza más espacios verdes y abiertos, terrazas, balcones más grandes, lugares dentro del hogar para tener privacidad en los momentos de trabajo y *home office* y, en muchos casos hay quienes escapan de las zonas más densas, concurridas y con muchos comercios”. El director de la firma Predial, Gabriel Brodsky, explica que previo a la pandemia, el valor del metro cuadrado de una terraza o patio se calculaba a un tercio del valor de uno cubierto, y actualmente esa proporción escaló al 50%. Incluso, si el espacio descubierto posee comodidades como parrilla, jacuzzi, o incluso paisajismo, puede valer un 60 o 70% de lo que sería la superficie cubierta.

Por otro lado, los espacios comunes en los edificios ganaron terreno después de que la cuarentena pusiera en evidencia la necesidad de compartir y convivir con el otro. El trabajo remoto, reforzado luego de la pandemia, también reforzó esta característica cada vez más demandada en departamentos. Más adelante, se analizará el peso de esta variable en inmuebles por sobre el peso de la cantidad de metros cuadrados o características del departamento en sí.

Mudafy asegura que “un 25% de los clientes le da una importancia especial a los amenities porque buscan habitar el edificio y no solamente el departamento”. Entre algunas de las amenities más demandadas se encuentran los gimnasios, espacios de trabajo compartido, piscina, espacios para mascotas, áreas de juego para niños, seguridad y salón o parrilla de usos múltiples, según indica el portal inmobiliario Urbania (2020).

Entre los nuevos hábitos adoptados al día de hoy, uno de los más interesantes resulta ser el teletrabajo. El modelo híbrido laboral surgió como una respuesta óptima ante el distanciamiento social, dejando atrás en muchos sectores e industrias, la dinámica presencial. Ante este nuevo escenario, la descentralización de las viviendas emergió como una respuesta natural para muchas personas que no necesitaban estar ubicados estratégicamente y buscaban mas tranquilidad.

Se abrió el abanico de opciones a todos aquellos que buscaban su próximo hogar, mas conscientes también de la necesidad de espacio para trabajar desde sus casas. Para muchas personas los hogares se transformaron definitivamente en su nuevo lugar de trabajo y estudio.

Con la transformación de las viviendas en centros multifuncionales de trabajo, ocio y educación, se ha intensificado la demanda de espacios de vida versátiles. Cada vez son más los inmuebles con oficinas en casa, diseños flexibles y espacios adaptables. Los promotores inmobiliarios están respondiendo con viviendas que satisfacen las diversas necesidades de la vida moderna. Un informe de Newmark (2023), consultora líder de bienes raíces, señala esta flexibilidad mediante “espacios de vivienda convertidos en oficinas, locales comerciales convertidos en centros de distribución, edificios convertidos en espacios de coworking” y otros ejemplos de cómo se representa el uso de inmuebles en respuesta a los cambios en la sociedad.

Muchas parejas de alrededor de 40 años hoy asumen que, incluso post-pandemia, seguirán trabajando muchas horas desde la casa, y eso los lleva a replantear la vivienda y los espacios que necesitan, explica Tomás Seeber, director de la inmobiliaria RG Montes.

Los amenities se readaptan y pasan a ser mucho más interesantes para el uso y funcionalidad en la vida diaria. También el comercio electrónico cambió mucho el modo de

comprar, con lo cual los edificios tienen que estar preparados para la recepción de paquetería (los llamados *pick up points*, que son los lugares adonde el propietario o inquilino puede ir a retirar los productos solicitados).

Por otro lado, la integración tecnológica es un elemento crucial en la lista de deseos del comprador de vivienda moderno, tal como indica Inmobiliarme (2024). El comprador de viviendas post-pandémico da prioridad a la integración de la tecnología en sus espacios vitales. Las características de las casas inteligentes, la infraestructura de Internet de alta velocidad y los servicios sin contacto se han convertido en aspectos integrales del valor de la propiedad. Los profesionales inmobiliarios se están adaptando a esta demanda destacando las características tecnológicas y asegurándose de que las propiedades estén equipadas para la era digital.

1.2. La ubicación como variable principal

Algunos años atrás, los barrios preferidos de Buenos Aires reflejaban diferentes prioridades en comparación con los actuales. En ese entonces, la ubicación céntrica era altamente apreciada, con una fuerte preferencia por los departamentos situados cerca de centros de negocios y áreas comerciales, como el microcentro y barrios emblemáticos como Palermo y Recoleta.

Posterior a la pandemia, distintas investigaciones comprueban lo que muchas personas tuvieron que atravesar al quedarse encerrados sin espacios descubiertos en sus casas, y alejados de zonas verdes. Un informe de la Organización Mundial de la Salud realizado en 2023 demuestra que pasar tiempo en la naturaleza (incluidas las zonas urbanas y periurbanas) mejora el estado de ánimo, la mentalidad y la salud mental. Las investigaciones demuestran que la exposición a parques y jardines puede incluso dejar atrás el estrés cotidiano. Y es así como las personas comenzaron a necesitar una mayor cercanía a espacios verdes, y a priorizar los espacios descubiertos dentro de sus casas, que quizás antes daban por sentado.

En el mismo sentido, las ubicaciones alejadas del centro de Capital Federal, en zonas más tranquilas o residenciales comenzaron a resultar más interesantes debido a su cercanía con espacios verdes y al menor valor que tenía el metro cuadrado en dichas zonas. Comenzó a

observarse una clara necesidad de irse a otros puntos de la ciudad o hallar propiedades con más espacio exterior.

La accesibilidad y la proximidad a estaciones de subte y colectivos también eran importantes, facilitando el transporte diario. Sin embargo, hoy en día es evidente que incluso la gente busca alejarse de lugares tan transitados, preservando su tranquilidad por sobre la comodidad. Barrios como Tigre y Escobar, donde el precio por metro cuadrado es más accesible y hay abundante espacio verde, están viendo un creciente interés. Cada vez más personas se mudan fuera de Capital Federal en busca de tranquilidad. Además, con el auge del trabajo remoto, la necesidad de ubicarse estratégicamente en el centro de la ciudad ha disminuido, permitiendo que los individuos prioricen otros aspectos en la elección de su hogar.

Vivir en la Ciudad de Buenos Aires tiene tanto beneficios como desventajas, y una de estas es el hecho de que buena parte de sus más de 3 millones de habitantes habita en departamentos sin espacios verdes al aire libre e incluso hasta sin balcón.

Puerto Madero ha sido históricamente un barrio con precios exageradamente más altos que el promedio. Si bien el barrio goza de prestigio y popularidad, cabe resaltar que la mayor parte de la vegetación natural de esta zona se concentra en dicho barrio, con especial énfasis en la Reserva Ecológica Costanera Sur. Este dato fue provisto por un informe de Properati, en el cual se propuso exponer los barrios con mayor cantidad de zonas verdes. Los barrios de Belgrano, Núñez, Colegiales y Palermo cuentan con un 80,33% de área urbana y poco menos de un 20% de vegetación.

Por otro lado, los barrios de Balvanera, San Cristóbal, Almagro, Boedo, Floresta, Monte Castro, Vélez Sarsfield, Versalles, Villa Luro, Villa Real, Villa General Mitre, Villa Devoto, Villa del Parque y Villa Santa Rita tienen un 100% de su territorio catalogado como "área urbana". El resto de los barrios poseen un área urbana de entre un 92,7% y un 98,7% de su territorio total. Se pretende analizar a continuación dichas observaciones reflejadas en los precios de venta.

Específicamente en el rubro inmobiliario, en la segunda mitad de la década de los noventa era habitual hablar de una media de 1000 USD por metro cuadrado de superficie para

unidades residenciales de departamentos en Capital Federal. Existiendo desde luego diferencias entre las cotizaciones de los distintos barrios integrantes de la ciudad, en función de la calidad del entorno, presencia de medios de transporte, oferta de equipamiento comercial, homogeneidad de tipologías constructivas y de nivel sociocultural de sus habitantes.

La zona norte del área urbana de la ciudad era la que contaba con los valores superiores. En el otro extremo, la zona sur, acusaba los registros más bajos de valuación. La brecha existente entre los valores reinantes en una y otra localización para unidades de similar calidad constructiva y comodidades difícilmente superaba el 33%.

Por ejemplo, el barrio Belgrano, uno de los más buscados por los porteños, tuvo grandes fluctuaciones en precio en los últimos años. Aunque éste sigue siendo un barrio prestigioso y con buena infraestructura, ya no goza del mismo nivel de codiciabilidad y demanda que tenía hace algunos años, debido a la combinación de ajustes en precios, variación en preferencias de los compradores, y un mercado inmobiliario más competitivo. Tal como indica ZonaProp (2021) luego de la pandemia por coronavirus, “aproximadamente el 32% de los departamentos en venta en Belgrano, han sido retasados a un precio menor”. Es por esto que estar al tanto de las tendencias del mercado y comprender cómo la ubicación influye en ellas es “fundamental para tomar decisiones informadas al comprar o vender una propiedad” (De Inmobiliarios, 2023).

Las casas en venta en barrios cerrados y countries (publicadas en precio de mercado) pasaron a tener gran aumento en la demanda. Hoy, el parque, la seguridad y los amenities son muy valorados.

1.3. La integración tecnológica en el mercado inmobiliario

La pandemia ha tenido un impacto significativo en el mercado inmobiliario y en los negocios relacionados. Desde los cambios en la demanda hasta la forma en que las empresas inmobiliarias operan.

La tecnología también ha revolucionado el mercado inmobiliario durante la pandemia. El uso de la inteligencia artificial (IA) ha permitido a las empresas inmobiliarias recopilar y analizar

datos con mayor precisión y eficiencia, lo que les ha ayudado a tomar decisiones más informadas en cuanto a precios, estrategias de marketing y otras cuestiones relacionadas con el mercado inmobiliario.

PropTech (Property Technology), también llamado *ReTech* (Real Estate Technology), es un término que combina los conceptos *property* y *technology*, y se refiere a empresas que utilizan tecnología para mejorar o reinventar servicios dentro del sector inmobiliario. Algunas de las tecnologías más destacadas utilizadas por el proptech incluyen el aprendizaje automático, *Big Data* y la geolocalización.

En términos generales, *PropTech* describe productos digitales para el sector de bienes raíces. *PropTech* se considera un sector en sí mismo pues, las disrupciones que inserta en los procesos tradicionales de la industria de bienes raíces se observan nuevos comportamientos, nuevos paradigmas y se produce una resignificación del mercado inmobiliario en sí mismo.

Así el mercado inmobiliario se integra a los mercados liderados por el análisis de datos, también con el sector *Fintech* (Tecnología y finanzas) y en general con todos los mercados que, atravesados por la tecnología, que abren el espectro de captura de datos sobre los comportamientos de consumo y las decisiones de las personas. La idea detrás de dicho sector es transformar la industria inmobiliaria mediante la integración de tecnología de vanguardia con *Big Data* y análisis predictivos para hacer que el trabajo de los profesionales sea más intuitivo y fácil de usar.

Plataformas como Zonaprop y Argenprop han empezado a incorporar estas funcionalidades, aunque aún de manera inicial. Los usuarios pueden ver tendencias de precios en distintos barrios, analizar la evolución histórica de los precios y obtener estimaciones de valor futuro. Zonaprop incluso posee un servicio de valuación de propiedades donde el usuario puede conocer el valor aproximado de una propiedad al ingresar las características del inmueble.

Portales como el de Properati también han desarrollado herramientas como el “preciómetro”, la cual en muy pocos clicks, permite conocer el valor de venta o alquiler de una propiedad. A pesar del considerable esfuerzo de muchas empresas por integrar tecnología, y

especialmente algoritmos de aprendizaje automático, en sus servicios, los resultados siguen siendo en gran medida preliminares.

2. Construcción y Análisis de la Base de Datos

El siguiente capítulo pretende dar cuenta de la metodología que se llevará a cabo para alcanzar los objetivos del trabajo, y respaldar las decisiones tomadas por otras investigaciones o trabajos previos que las respalden. El propósito es demostrar que la propuesta presente está alineada con el conocimiento existente y tiene sustento en investigaciones previas reconocidas en la disciplina.

2.1. Selección de Datos

Para alcanzar los objetivos propuestos con los enfoques mencionados, se utilizaron dos bases de datos que logren representar los escenarios pre y post pandemia. Se recopilaron datos históricos de propiedades en venta en Buenos Aires durante los años 2020 a 2024.

El conjunto de datos utilizado consiste en dos bases de datos distintas, la primera abarcando el período 2020-2021 y la segunda 2023-2024. Ambas contienen información histórica sobre propiedades vendidas, incluyendo datos como los precios de venta, ubicación, tamaño y características de los departamentos.

Los datos correspondientes al período 2020-2021 se obtuvieron de la plataforma Properati, la cual ofrece el acceso libre a este tipo de información. Esta base contiene 1.000.000 de observaciones que incluyen publicaciones de venta y de alquiler de distintos tipos de inmuebles en Buenos Aires. Para realizar un análisis específico y eficiente, el enfoque del trabajo fue dirigido a los departamentos en venta de 1, 2 y 3 ambientes en Capital Federal. El resultado final de observaciones de la muestra terminó en 60.000 registros.

Por otro lado, los datos correspondientes al período 2023-2024 fueron conseguidos mediante la metodología *web scraping*, de portales como ArgenProp, ZonaProp y Miranda Bosch. *Web scraping* refiere al proceso de recolectar datos desde la web, ya sea de forma manual o automática. Esta técnica extrae el contenido HTML de páginas web para filtrar la información

requerida y almacenarla, de forma comparable al proceso automático de copiado y pegado. Python es considerado uno de los mejores lenguajes de programación para realizar tareas de *web scraping* debido a su velocidad y facilidad de uso. Específicamente para el presente trabajo se utilizaron las librerías Selenium y BeautifulSoup.

Se enfrentaron grandes desafíos ya que cada vez más, los sitios web protegen su información online con bloqueos anti scraping, por lo que fue necesario contar con cierta familiarización con técnicas avanzadas. Como resultado final, se obtuvo una base de 8.000 registros incluyendo departamentos y casas en alquiler y venta. Para realizar un análisis equitativo y comparar ambas bases con características similares, se filtraron los departamentos en venta de 1, 2 y 3 ambientes en Capital Federal, resultando así en una base de 2.200 registros.

Para una mejor identificación en el trabajo, se referirá a la base de 2020-2021 como el conjunto de datos pre-pandemia, mientras que el set de datos 2023-2024 será nombrado datos post-pandemia.

2.2. Limpieza de Datos

Inicialmente se llevó a cabo una exhaustiva limpieza de datos, ya que las bases de datos fueron generadas manualmente por distintos usuarios y por ende contenían una gran cantidad de errores que, de no ser tratados correctamente, podrían llevar a conclusiones erróneas y un alto sesgo.

La limpieza de datos consistió entonces, en reconocer valores erróneos, extremos o *outliers*, imputar valores faltantes mediante distintas técnicas como centroide, el vecino más cercano o por la mediana. Mediante la técnica de *text mining* se crearon campos que den cuenta de variables valiosas, como ser un departamento a estrenar, con piscina o cochera. Dichos atributos fueron convertidos en variables cuantitativas de tipo booleano para comprender el peso que influyen en los respectivos precios.

En cuanto a la segunda base, ya que la misma fue extraída mediante la técnica de *web scraping*, el cuidado de manipulación de datos debió ser extremo para evitar cualquier tipo de

error que pudiera llevar a conclusiones erróneas. Se utilizaron gráficos de visualización como *boxplot* para identificar posibles errores de extracción de datos.

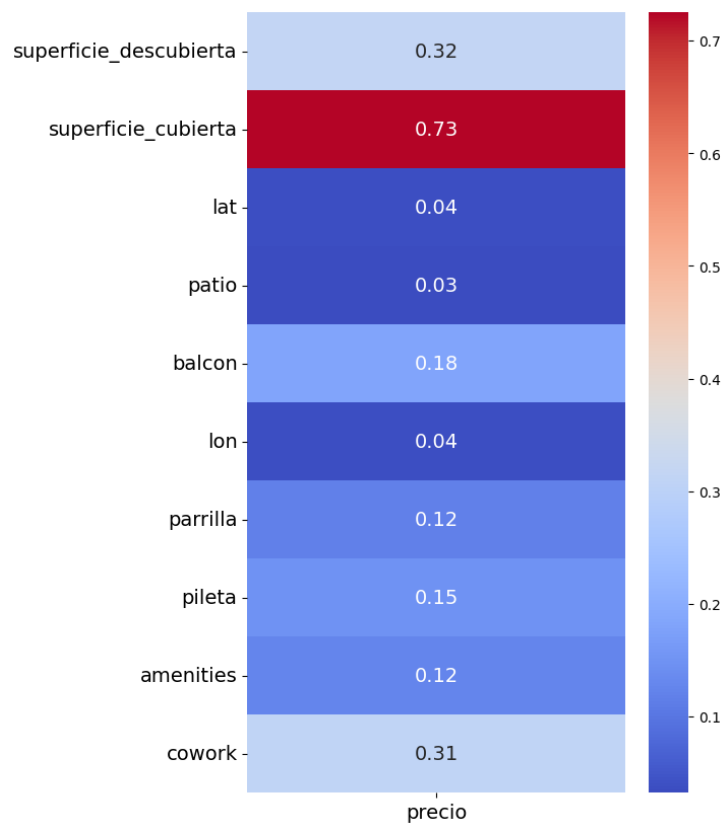
Luego de realizar la pertinente limpieza, las variables seleccionadas para proceder a realizar el análisis exploratorio fueron: barrio, latitud, longitud, superficie cubierta, superficie descubierta, parrilla, balcón, amenities, cowork, cochera y precio. Se creó una variable de precio por metro cuadrado, dividiendo el valor del precio por la superficie cubierta, y se agrupó la misma por barrio.

Si bien como fue mencionado anteriormente, la segunda base es mucho más pequeña que la primera, se logró alcanzar estándares de extrema exactitud para la misma, logrando así una muestra fiel y representativa de la realidad. Una vez realizada la correspondiente limpieza de datos, se procedió al inicio de la estadística descriptiva y el análisis exploratorio.

2.3. Análisis Exploratorio

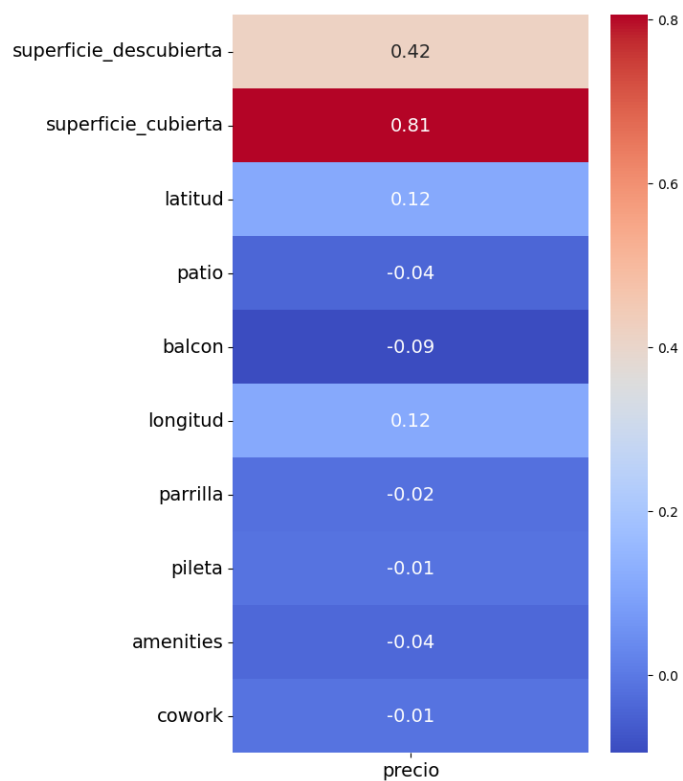
Uno de los objetivos específicos mencionados inicialmente fue comprobar que la superficie descubierta de los inmuebles se revalorizó y pasó a cobrar mayor importancia después de la pandemia. Para comprobar esta hipótesis, se utilizaron dos matrices de correlación que miden la relación entre distintas variables (como precio, superficie cubierta, superficie descubierta, y amenities) en los dos periodos de tiempo analizados: 2020-2021 (pre-pandemia e inicios de la pandemia) y 2023-2024 (post-pandemia).

Gráfico 1: Correlación de Precio contra las Principales Variables 2020-2021



Fuente: Elaboración propia. Período 2020-2021.

Gráfico 2: Correlación de Precio contra las Principales Variables 2023-2024



Fuente: Elaboración propia. Período 2023-2024.

A modo de ilustración, se exponen los resultados en una tabla a continuación para una mejor lectura de los datos:

Tabla 1: Correlación de Precio contra las Principales Variables en ambos períodos

Variabes	Índice de Correlación 2020-2021	Índice de Correlación 2023-2024
Superficie descubierta	0.32	0.42
Superficie cubierta	0.73	0.81
Patio	0.03	-0.04
Balcón	0.18	-0.09
Parrilla	0.12	-0.02
Pileta	0.15	-0.01
Amenities	0.12	-0.04
Espacio de <i>coworking</i>	0.31	-0.01
Latitud	0.04	0.12
Longitud	0.04	0.12

Fuente: Elaboración propia. Período 2020-2024.

Es posible observar que en el primer período, la correlación entre el precio y la superficie descubierta era de 0.32, indicando una relación moderada. En el período post-pandemia, esta correlación (precio y superficie descubierta) aumentó a 0.42, mostrando un leve cambio en la importancia que se le otorga a los espacios al aire libre en la valoración de las viviendas.

El aumento en la correlación sugiere que, a raíz de la pandemia, los espacios descubiertos han pasado a ser elementos clave en la valoración de una propiedad, lo cual es coherente con la creciente necesidad de disponer de espacios abiertos debido a las restricciones de movilidad y el tiempo prolongado en casa.

Para una mejor interpretación, se crearon variables para identificar los distintos tipos de espacio al aire libre que tienen las viviendas, y categorizar las principales *amenities* que poseen los departamentos. Al observar la evolución de la correlación de estas variables, no es posible confirmar la hipótesis de que, por ejemplo, un espacio de *coworking* o una pileta en un edificio aumente el precio de un departamento.

La pandemia sin duda ha impulsado la demanda de viviendas que ofrecen mayores comodidades para la vida diaria y el trabajo remoto, pero desafortunadamente no ha sido posible corroborarlo en la presente investigación. La razón principal se debe a que la construcción de dichas variables (amenities, cochera, pileta, cowork y balcón) fue realizada mediante la técnica de *text mining*, extrayendo la información existente en la descripción de cada departamento. Sin embargo, no todos los departamentos poseen información correcta o completa, por lo que resulta difícil mitigar los sesgos. La superficie cubierta mantuvo su alta correlación en el precio, sin experimentar un cambio drástico como la superficie descubierta.

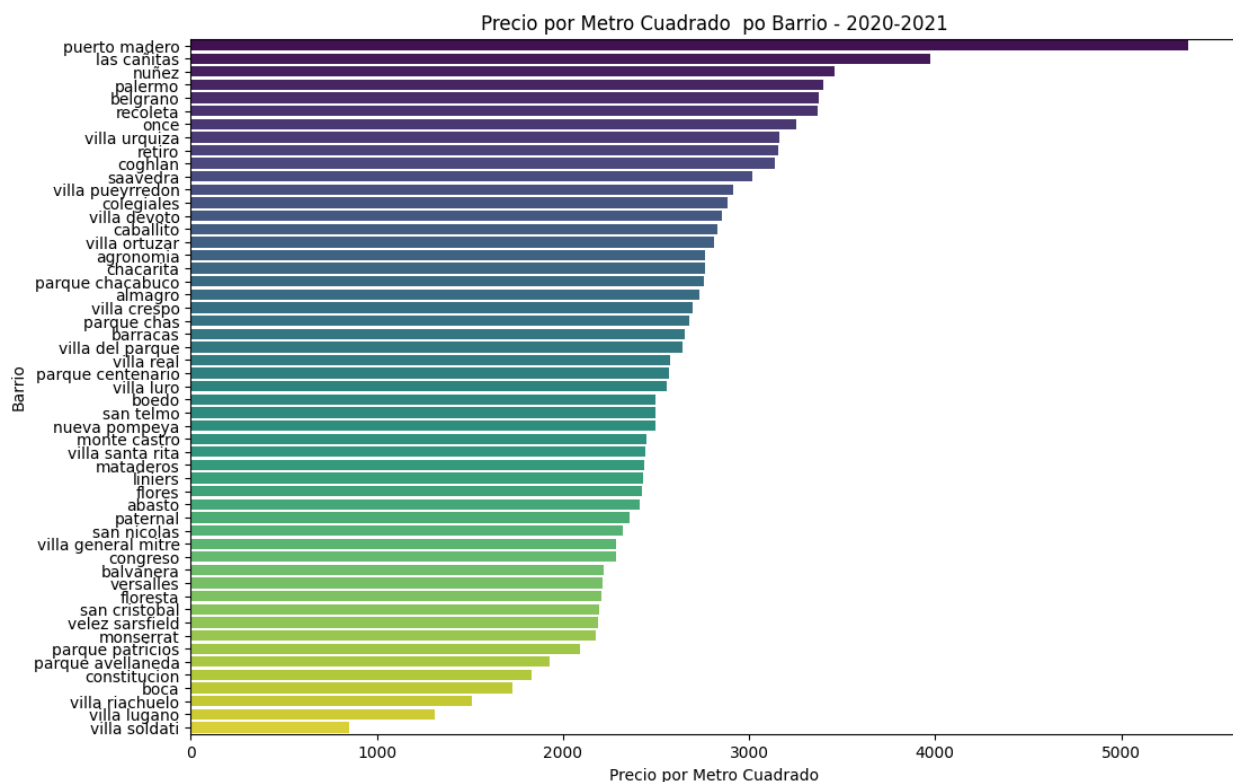
El proceso de *web scraping*, aunque útil para recopilar datos cuando no hay otras fuentes disponibles, puede ser propenso a errores de extracción debido a la falta de estandarización en las plataformas de origen. Esto incluye la posibilidad de datos incompletos o mal estructurados, lo que podría haber afectado la precisión del análisis, especialmente en el contexto post-pandemia, donde la cantidad de registros ya era limitada. Cualquier inconsistencia o error en los datos extraídos podría haber causado sesgos en el análisis.

Se llevo a cabo un análisis enfocado en comprender el valor de los metros cuadrados cubiertos en distintos barrios. Según las hipótesis iniciales, algunos barrios podrían haber ganado valor por sus ubicaciones próximas a espacios verdes. Como se mencionó en los inicios del trabajo, Puerto Madero ha sido históricamente un barrio con precios altos y la mayor parte del mismo posee extensos espacios de vegetación natural. Es posible ver en los gráficos a continuación, que el valor del metro cuadrado en Puerto Madero aumentó significativamente de un período a otro, siendo el barrio con mayor aumento de valores.

Es posible observar que muchos barrios bajaron sus valores, y la razón de esto yace en explicaciones económicas y políticas que atravesó la sociedad. Sin embargo, es posible ver que las reducciones más bajas estuvieron en los barrios como Nuñez y Palermo, registrando disminuciones de 9% y 1%, respectivamente. Colegiales, otro barrio destacado por sus espacios verdes, tuvo un aumento de valor (+34%). Belgrano y Recoleta, barrios anteriormente destacados por sus ubicaciones estratégicas y populares, bajaron significativamente el valor, con caídas del 13% y 17%, respectivamente. .

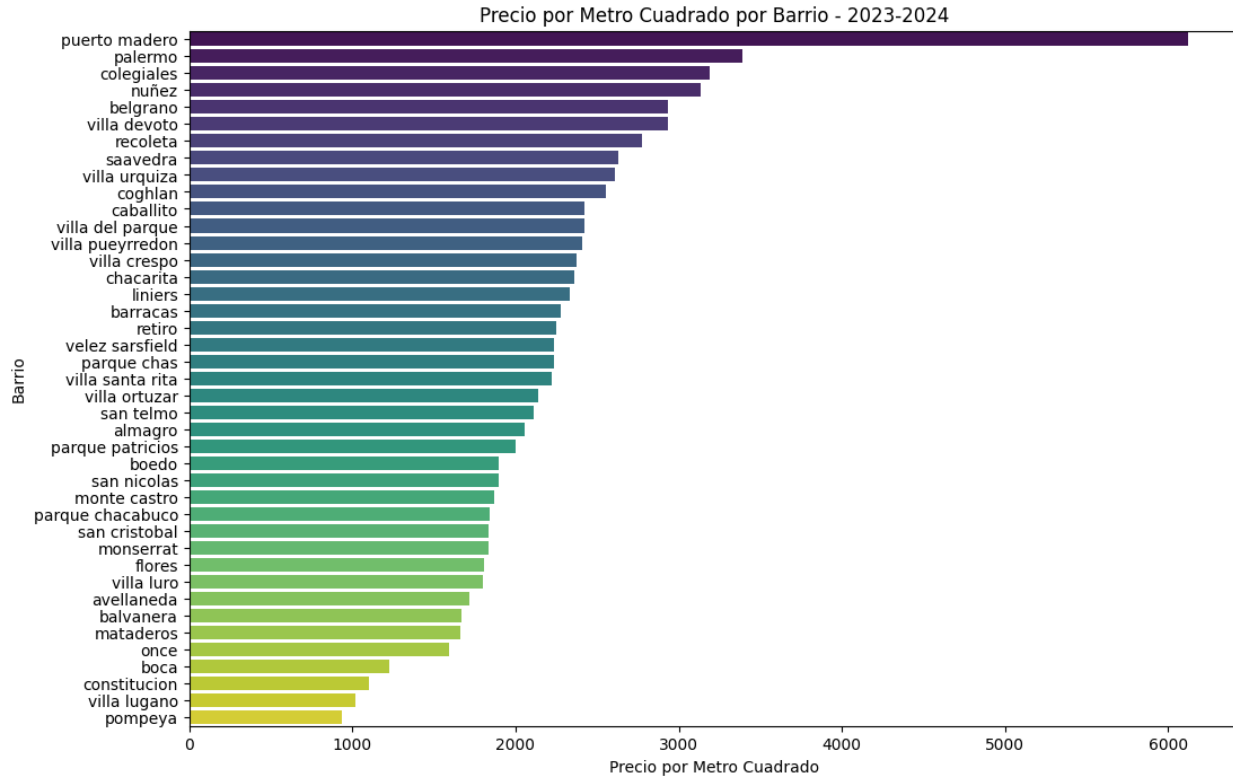
Es posible observar que hoy en día, los ciudadanos buscan zonas grandes con pulmones verdes. En este sentido, Villa Devoto experimentó un salto de precios, de alrededor de un 9%, reflejado en su característica de poseer más del 30% de cobertura de arbolado urbano.

Gráfico 3: Precio por m2 por barrio - 2020-2021



Fuente: Elaboración propia. Período 2020-2021.

Gráfico 4: Precio por m2 por barrio - 2023-2024



Fuente: Elaboración propia. Período 2023-2024

3. El Rol de los Modelos Predictivos en las Plataformas Digitales

Según Molinari (2020), la integración de modelos predictivos en plataformas digitales del mercado inmobiliario en Argentina ha comenzado a transformar la industria. Si bien su existencia es incipiente y su accesibilidad acotada, han demostrado mejorar la toma de decisiones para compradores y vendedores que han declarado utilizarlos. Estas plataformas no solo muestran las propiedades disponibles, sino que también proporcionan análisis y proyecciones de precios basados en datos predictivos. Dicha integración ofrece a los usuarios una herramienta poderosa para tomar decisiones más informadas y estratégicas.

Se aplicaron técnicas de análisis estadístico y modelos de aprendizaje automático para desarrollar el modelo predictivo en la plataforma de Google Colab, con el lenguaje Python. Se

seleccionaron determinados algoritmos y se evaluó sus respectivos rendimientos, comparando los resultados.

3.1. Desarrollo de modelos de predicción

Para evaluar los distintos modelos de predicción y elegir uno entre ellos, se prestó atención a la distancia de sus predicciones del precio con los valores reales. La diferencia entre la predicción y el valor real es el error. Se eligió en base a esta métrica, utilizar dos modelos distintos, el primero llamado Random Forest y el segundo Gradient Boosting.

El modelo Random Forest es un algoritmo de aprendizaje supervisado que funciona construyendo múltiples árboles de decisión en el proceso de entrenamiento y generando la predicción final promediando los resultados de cada uno de estos árboles. Es robusto frente a la variabilidad en los datos y suele tener un buen desempeño en una amplia variedad de problemas, incluida la predicción de precios. Además, Random Forest es eficaz para evitar el sobreajuste (*overfitting*), ya que combina las predicciones de múltiples árboles para obtener un resultado más generalizado.

Gradient Boosting es un método de aprendizaje supervisado que construye modelos de predicción de manera secuencial, lo que significa que cada nuevo modelo corrige los errores de los modelos anteriores. Esto se realiza ajustando pequeños modelos débiles (normalmente árboles de decisión poco profundos) de manera iterativa, con el objetivo de reducir el error residual del conjunto anterior.

Se analizaron distintas combinaciones de hiperparámetros en ambos modelos, evaluando siempre el RMSE (error medio cuadrático). RMSE se puede interpretar como la desviación estándar de la varianza inexplicada, y tiene la propiedad útil de estar en las mismas unidades que la variable de respuesta. Los valores más bajos de RMSE indican un mejor ajuste. RMSE es una buena medida de la precisión con que el modelo predice la respuesta, y es el criterio más importante para ajustar si el propósito principal del modelo es la predicción.

Se probaron distintas combinaciones de parámetros para ambos modelos, concluyendo finalmente en $n_estimators=1000$, $learning_rate=0.1$, $max_depth=9$. El primer hiperparámetro ($n_estimators$) define el número de árboles individuales que el modelo construirá. En el contexto de Gradient Boosting, donde el modelo se construye de manera secuencial, cada árbol trata de corregir los errores del árbol anterior. Un valor de 1000 significa que se construirán 1000 árboles. Un número mayor de estimadores puede hacer que el modelo sea más potente, pero también aumenta el riesgo de sobreajuste (*overfitting*) si no se ajustan bien otros hiperparámetros, como el $learning_rate$.

El *learning rate* es la tasa de aprendizaje, la cual controla la contribución de cada árbol en la corrección de errores. Esencialmente, es un factor de escalado aplicado a las predicciones de cada árbol antes de sumarlas al modelo. Una tasa de aprendizaje de 0.1 significa que cada árbol contribuye solo un 10% al modelo final.

Por último, max_depth define la profundidad máxima de cada árbol individual. Este hiperparámetro controla cuántas divisiones o nodos puede tener cada árbol, afectando la complejidad del árbol. Un valor de 9 indica que los árboles pueden tener hasta 9 niveles de profundidad, lo que les permite capturar relaciones complejas en los datos. Sin embargo, árboles más profundos pueden llevar a un sobreajuste, donde el modelo se ajusta demasiado a las peculiaridades del conjunto de entrenamiento en lugar de generalizar bien a nuevos datos.

3.2. Optimización y Resultados

Random Forest fue descartado por mostrar un mayor error cuadrático medio (RMSE) en comparación con Gradient Boosting. Esto se debe a que, si bien es eficaz en escenarios generales, no pudo capturar de manera óptima las complejas interacciones entre las variables en este conjunto de datos, especialmente aquellas que han ganado relevancia en el contexto post-pandemia.

Por otro lado, Gradient Boosting, al ser un modelo secuencial, permitió ajustar mejor las predicciones mediante la corrección iterativa de los errores, logrando un mejor rendimiento en este caso. El modelo de predicción desarrollado alcanzó un error medio cuadrático de 19.307,

para la base pre-pandemia. Este resultado indica una buena capacidad del modelo para aproximar los valores reales de las propiedades en el mercado, lo que lo hace una herramienta valiosa para la toma de decisiones dentro del sector inmobiliario.

En términos de confiabilidad, el modelo puede considerarse sólido debido a la metodología rigurosa empleada en su desarrollo. La elección cuidadosa de los hiperparámetros y la validación a través del RMSE garantizan que el modelo no solo es preciso, sino que también es capaz de generalizar bien en datos no vistos. Sin embargo, como con cualquier modelo predictivo, es esencial que Buenos Aires Real State o cualquiera sea la empresa que lo incorpore continúe monitoreando su rendimiento y lo actualice regularmente para adaptarse a cambios en las condiciones del mercado.

Para el periodo post-pandemia, el RMSE fue mayor (140.000), lo que sugiere que aún se pueden realizar mejoras en el ajuste de hiperparámetros o en la inclusión de nuevas variables que reflejen mejor los cambios en el mercado. La optimización continua del modelo es fundamental, especialmente en un mercado tan dinámico como el inmobiliario, que ha experimentado transformaciones estructurales tras la pandemia. El mayor RMSE post-pandemia refleja las dificultades del modelo para capturar las dinámicas emergentes del mercado tras los cambios estructurales que trajo la pandemia. Esto sugiere que podrían ser necesarias variables adicionales o ajustes en los hiperparámetros para mejorar la precisión en este nuevo contexto.

Se decidió excluir el barrio Puerto Madero de la base post-pandemia, para evaluar mejoras para reducir el error de predicción. Sin embargo, dicha medida no logró optimizar significativamente el RMSE del modelo post-pandemia (tan sólo un 4%).

Para garantizar la solidez del modelo predictivo, se utilizó la validación cruzada en 10 partes como técnica de evaluación del rendimiento. Esta técnica permite dividir el conjunto de datos en diez subconjuntos o *folds* de tamaño aproximadamente igual. Durante cada iteración, el modelo se entrena con 9 de los subconjuntos y se prueba en el subconjunto restante. Este proceso se repite diez veces, utilizando cada subconjunto como conjunto de prueba en una ocasión.

distinta, lo que asegura que cada punto de datos se utilice tanto para entrenamiento como para validación.

Al finalizar las 10 iteraciones, se calcula el promedio de los RMSE obtenidos en cada uno de los *folds*. Este valor promedio proporciona una estimación robusta de la capacidad del modelo para generalizar a nuevos datos. Además, se calculó la desviación estándar de los errores para evaluar la variabilidad del rendimiento del modelo.

La validación cruzada ayuda a minimizar el riesgo de sobreajuste al garantizar que el modelo no se ajuste excesivamente a un único subconjunto de datos, lo que permite medir su capacidad de generalización de manera más precisa. En este caso, el uso de un aprendizaje secuencial en Gradient Boosting se complementó con esta técnica para asegurarse de que el modelo no estuviera ajustado excesivamente a un subconjunto específico de datos.

Tras la validación cruzada en 10 partes, el resultado de RMSE promedio fue de 20.377, y la desviación estándar del RMSE fue de 840. Estos resultados indican que, en promedio, el modelo Gradient Boosting tiene un error de predicción de aproximadamente 20.377 USD. Este valor es una medida de la precisión del modelo en predecir los precios de las propiedades, y la baja desviación estándar sugiere que el modelo es consistente en su rendimiento a través de los diferentes subconjuntos de datos.

A pesar de los esfuerzos realizados para limpiar y preparar los datos, es importante destacar cómo la calidad de los datos y el pequeño tamaño de la muestra post-pandemia podrían haber afectado los resultados del análisis. Mientras que la base de datos pre-pandemia contenía 60,000 registros, la base post-pandemia se redujo significativamente a 2.200 registros. Esta diferencia en el tamaño de la muestra puede haber influido en la capacidad del modelo para capturar de manera precisa las nuevas tendencias del mercado inmobiliario post-pandemia. Una muestra más pequeña introduce mayor variabilidad y puede hacer que el modelo sea más sensible a fluctuaciones o outliers, lo que podría explicar el mayor RMSE observado en este período. Además, una muestra limitada reduce la capacidad del modelo para identificar patrones

generales, ya que es más probable que las relaciones detectadas estén influenciadas por un subconjunto específico de propiedades.

Por último en cuanto a los resultados del modelo de predicción, se observó la importancia de las distintas variables más relevantes que explican el precio de las propiedades. En el caso del modelo pre-pandemia, la superficie cubierta tuvo una importancia del 45%, siendo por mucho, el factor más determinante en la predicción del precio de las propiedades antes de la pandemia. Esto refleja la tendencia de los compradores de valorar principalmente el espacio cubierto como una de las principales métricas para determinar el valor de un inmueble. Otras características menores como coworking, cochera, y amenities tienen una relevancia marginal (entre 1% y 2%), lo que indica que, antes de la pandemia, estos aspectos no eran tan valorados por los compradores.

En el modelo post-pandemia se observó un cambio significativo en las variables más importantes. La superficie cubierta sigue siendo la variable más importante con un peso del 84.8%, aunque su relevancia ha aumentado considerablemente en comparación con el modelo pre-pandemia. Esto sugiere que, tras la pandemia, los compradores continúan valorando fuertemente el espacio cubierto, posiblemente porque el confinamiento destacó la importancia de áreas amplias y cómodas dentro del hogar.

Una de las diferencias más notables es el aumento en la importancia de la superficie descubierta, que pasa a ocupar el segundo lugar en relevancia, reflejando cómo la pandemia ha cambiado las preferencias, haciendo que los espacios abiertos, como patios o terrazas, sean mucho más valorados debido a la necesidad de disfrutar de espacios al aire libre.

La importancia de la cochera también aumenta lo que indica que la disponibilidad de estacionamiento ha ganado relevancia, quizás porque la pandemia aumentó el uso del transporte privado.

3.3. Aplicación de modelos de predicción

Buenos Aires Real State se va a ver altamente beneficiada de la utilización del modelo de predicción. La integración de este tipo de herramientas, y específicamente de este modelo predictivo puede ofrecer múltiples beneficios e incluso crear nuevas oportunidades de negocio.

En primer lugar, proporciona a la empresa una ventaja competitiva al ofrecer a sus usuarios (tanto compradores como vendedores) una herramienta avanzada que les permita tomar decisiones más informadas. Esto puede traducirse en una mayor satisfacción del cliente, una reducción en el tiempo de venta de las propiedades y, potencialmente, en un aumento de la cuota de mercado.

Además, al utilizar un modelo confiable y bien calibrado, la empresa puede mejorar la eficiencia de sus operaciones internas, optimizando precios y estrategias de marketing con mayor precisión. Esto no solo puede conducir a una mejor alineación entre la oferta y la demanda, sino también a una mayor rentabilidad a largo plazo.

En resumen, la adopción de este modelo de predicción no solo es técnicamente viable, sino también estratégicamente beneficiosa, proporcionando tanto a la empresa como a sus clientes herramientas valiosas para navegar en un mercado inmobiliario cada vez más competitivo.

Conclusiones

El análisis del impacto de la pandemia en el mercado inmobiliario de Buenos Aires ha revelado cambios significativos en las prioridades y necesidades de los compradores, lo que ha afectado la valoración de las propiedades. La investigación ha demostrado que variables que anteriormente eran secundarias, como los espacios descubiertos y la cercanía a áreas verdes, han ganado una relevancia considerable en la determinación del precio de los inmuebles post-pandemia.

La investigación ha identificado un aumento en la correlación entre la superficie descubierta y el precio de los inmuebles, pasando de 0.32 en el periodo 2020-2021 a 0.42 en 2023-2024. Sin embargo, no se observó un incremento significativo en el valor de los amenities, como patios y piletas, lo cual se lo atribuye a una necesidad de mejora en la manipulación de los datos, específicamente los de carácter de texto.

Los resultados del modelo predictivo, desarrollado utilizando técnicas de aprendizaje automático, han mostrado una mejora en la precisión de las predicciones de precios, con un RMSE de aproximadamente 19.000. Esto indica que el modelo es capaz de captar de manera adecuada las nuevas tendencias y variables relevantes en el mercado inmobiliario, con un error de predicción de alrededor 19.000 USD.

Próximos Pasos

- **Desarrollo de Variables:** Dado que Puerto Madero y otros barrios con comportamientos atípicos afectan negativamente el ajuste general del modelo, podría haber variables clave, específicas a estos barrios, que expliquen mejor las dinámicas de precios.
- **Ampliación de la Muestra:** Para mejorar la robustez y generalización del modelo, es crucial obtener una muestra más extensa y representativa que incluya datos de una mayor variedad de propiedades y periodos. Esto contribuirá a un modelo más sólido y adaptable a diferentes escenarios del mercado. Además, sería positivo para las hipótesis inicialmente mencionadas, incluir barrios periféricos de la Capital Federal para comprobar el incremento de la demanda en dichos barrios.

- **Exploración de Modelos más Avanzados:** Aunque el modelo actual ha mostrado buenos resultados, es importante explorar el uso de modelos más avanzados, como redes neuronales, que podrían capturar relaciones más complejas en los datos y potencialmente mejorar la precisión de las predicciones.

El informe destaca la importancia de seguir adaptando las herramientas predictivas a las nuevas realidades del mercado, utilizando bases de datos limpias y representativas, y considerando la inclusión de tecnologías avanzadas como el aprendizaje automático para mejorar la precisión de las estimaciones. En conjunto, estas conclusiones subrayan la necesidad de que las plataformas inmobiliarias y las empresas del sector sigan innovando y adaptándose a las nuevas demandas del mercado post-pandemia, asegurando así una ventaja competitiva y un mejor servicio a sus clientes.

Referencias bibliográficas

Hu, Y., Li, S., & Liu, Y. (2020). Machine learning approaches for predicting real estate prices. Journal of Real Estate Research.

Glaeser, E. L., Gyourko, J., & Saks, R. E. (2018). Urban growth and housing supply. Journal of Economic Geography.

Zhang, D., Zhou, L., & Zheng, C. (2019). Comparison of machine learning algorithms for predicting real estate prices. IEEE Transactions on Knowledge and Data Engineering, 31(2), 233-242.

Infobae. (2023, 13 de enero). Argentina tiene la segunda inflación más alta del mundo. Recuperado de <https://www.infobae.com/economia/2023/01/13/argentina-tiene-la-segunda-inflacion-mas-alta-de-l-mundo/>

Zonaprop. (2024). Informe de precios. Recuperado de <https://www.zonaprop.com.ar/informes-de-precios/>

Banco Central de la República Argentina. (2023). Informe de estabilidad financiera. Recuperado de https://www.bcra.gob.ar/Pdfs/PublicacionesEstadisticas/IEF_2023.pdf

Instituto Nacional de Estadística y Censos. (2023). Acceso a la vivienda y condiciones de los hogares. Recuperado de <https://www.indec.gob.ar/indec/web/Nivel4-Tema-4-53-84>

Forbes Digital. (2022). Buenos Aires es la tercera ciudad más cara de Sudamérica para comprar un inmueble. Recuperado de <https://www.forbesargentina.com/rankings/buenos-aires-tercera-ciudad-mas-cara-sudamerica-comprar-inmueble-n15683>

ZonaProp. (2021). Capital Federal: Estos son los precios de los alquileres barrio por barrio. Recuperado de

<https://www.informeconstruccion.com/nota/indices-e-informes/7203/capital-federal-estos-son-precios-alquileres-barrio-barrio.html>

De Inmobiliarios. (2023). ¿Cuál es el valor del m2 en Belgrano para la venta 2023? Recuperado de

<https://www.deinmobiliarios.com/blog/cual-es-el-valor-del-m2-en-belgrano-para-la-venta-2023>

La Nación. (2022). Amenities curiosos: Los nuevos espacios comunes que piden los propietarios en los edificios. Recuperado de

<https://www.lanacion.com.ar/propiedades/casas-y-departamentos/amenities-curiosos-los-nuevos-espacios-comunitarios-que-piden-los-propietarios-en-los-edificios-tras-nid11072022/>

Urbania. (2021). ¿Cuáles son los amenities más buscados por las personas? Recuperado de <https://urbania.pe/blog/noticia/cuales-son-los-amenities-mas-buscados-por-las-personas/>

De Inmobiliarios. (2024). ¿Qué se viene en el mercado inmobiliario en el 2024? Recuperado de <https://www.deinmobiliarios.com/blog/que-se-viene-en-el-mercado-inmobiliario-en-el-2024>

Reporte Inmobiliario. (2020). Crisis y precios inmobiliarios. Recuperado de <https://www.reporteinmobiliario.com/article3842-crisis-y-precios-inmobiliarios.html>

Inmobiliariame. (2024). El mercado inmobiliario post-pandemia. Recuperado de <https://inmobiliariame.com/mercado-inmobiliario-post-pandemia/>

National Geographic. (2023). ¿Por qué la naturaleza beneficia a la salud? Explicación científica. Recuperado de <https://www.nationalgeographic.es/medio-ambiente/2023/07/por-que-naturaleza-beneficios-salud-explicacion-cientifica>

Clarín. (2020). Balcones, patios y terrazas: La nueva obsesión de quienes buscan vivienda en plena cuarentena. Recuperado de https://www.clarin.com/sociedad/balcones-patios-terrazas-nueva-obsesion-buscan-vivienda-plena-cuarentena_0_HrqCkuI0u.html

Mudafy. (2021). Cómo se reconfiguró el mercado inmobiliario argentino en la pospandemia. Recuperado de <https://mudafy.com.ar/blog/post/como-se-reconfiguro-el-mercado-inmobiliario-argentino-en-la-pospandemia>

Cronista. (2023). CABA: ¿Cuáles son los barrios más verdes y con mejores precios para comprar propiedades? Recuperado de <https://www.cronista.com/negocios/caba-cuales-son-los-barrios-mas-verdes-y-con-mejores-precios-para-comprar-propiedades/>

Apéndices

Apéndice I - Conjunto de Datos utilizado para el período pre-pandemia 2020-2021:

Este apéndice presenta el conjunto de datos recopilados de la plataforma Properati para el período pre-pandemia. Contiene un total de 60.000 registros, incluyendo variables como precio, superficie cubierta, superficie descubierta y barrio.

Apéndice II - Conjunto de Datos utilizado para el período post-pandemia 2023-2024:

Este apéndice muestra el conjunto de datos obtenido a través de *web scraping* de portales inmobiliarios para el período post-pandemia, con un total de 2.238 registros. Se incluye información clave sobre las propiedades en venta, como ubicación y superficie descubierta.

Apéndice III - Cuaderno Colab con la limpieza de datos y armado de modelo del período pre-pandemia (2020-2021):

Este apéndice incluye un cuaderno de trabajo desarrollado en Google Colab que detalla el proceso de limpieza de los datos y la construcción del modelo predictivo utilizando información del mercado inmobiliario correspondiente al período pre-pandemia, abarcando los años 2020-2021. Se describen las técnicas aplicadas para la imputación de valores faltantes, la normalización de variables y los procedimientos de selección de características. Además, se documenta la metodología utilizada para entrenar un modelo de aprendizaje automático sobre los datos de este período, con el objetivo de analizar el comportamiento del mercado antes de que los efectos de la pandemia comenzaran a influir en las variables clave.

Apéndice IV - Cuaderno Colab con la limpieza de datos y armado de modelo del período post-pandemia (2023-2024):

Este apéndice presenta un cuaderno de trabajo en Google Colab, similar al Apéndice III, pero centrado en el período post-pandemia (2023-2024). Se sigue un proceso paralelo de limpieza de datos y construcción del modelo predictivo, ajustando los procedimientos a las características de los datos más recientes. El análisis se enfoca en las variables que han ganado relevancia en el contexto post-pandemia, como la presencia de espacios verdes y áreas

descubiertas, utilizando técnicas de aprendizaje automático para evaluar cómo estos factores influyen en los precios actuales del mercado inmobiliario.

Reporte del Mentor

El presente trabajo plantea el problema de analizar el impacto que tuvo la pandemia de Covid19 en el mercado inmobiliario de la ciudad de Buenos Aires. Considero que dicho problema es muy relevante en el contexto del mercado de compra y venta de inmuebles.

El problema está identificado y definido correctamente. Se realiza una descripción adecuada de las condiciones actuales y qué se pretende resolver con este trabajo. El interrogante es consistente con la problemática planteada. La hipótesis planteada de manera preliminar se definirá a medida que profundice el estado del conocimiento acerca del tema.

El planteo del problema, los objetivos y la hipótesis se encuentran articulados con la formación de base, la carrera profesional del autor y con la especialización que está realizando. Los conceptos abordados se encuentran dentro de los contenidos de los módulos de la especialización.

Además, considero que el objetivo general propuesto de identificar las variables con mayor impacto en el precio de los inmuebles en Buenos Aires en la actualidad, comprender su incidencia y, posteriormente, predecir los precios en función de estas variables. es coherente con el problema planteado. Y dicho objetivo es consistente con la hipótesis de que ciertas variables -anteriormente secundarias- han cobrado mayor protagonismo y hoy en día influyen más en el precio de los inmuebles que hace tres años.

Finalmente, considero que el presente trabajo alcanza el objetivo propuesto al presentar un detallado análisis de los datos y un correcto desarrollo del modelo de predicción.

Roberto Abalde