

# Detectar la transparencia de la información en el Mercado inmobiliario italiano: un enfoque de aprendizaje automático

Laura Gabrielli\*, Aurora Greta Ruggeri\*\*,  
Massimiliano Scarpa\*\*\*

Palabras clave: Análisis del mercado inmobiliario;  
Valor de mercado; Precio inicial; Transparencia del mercado;  
Aprendizaje automático; Redes neuronales artificiales

## Abstracto

Esta investigación tiene como objetivo comprender cómo la transparencia del mercado y la confiabilidad de los datos pueden influir en los procedimientos de valoración y los procesos de toma de decisiones en el mercado inmobiliario italiano.

A través del análisis de tres mercados inmobiliarios diferentes y la validación de la información recopilada, el objetivo de este artículo es comprender si, y en qué medida, el uso de precios de venta en lugar de precios reales de compra y venta puede conducir a errores de valoración, aumentar la incertidumbre de valoración y socavar los procesos de toma de decisiones de inversión.

Los resultados de la investigación destacan las principales fuentes de opacidad de la información en el mercado inmobiliario italiano, clasificándolas según su impacto en el valor inmobiliario.

La novedad de esta investigación radica en el uso integrado de técnicas de aprendizaje automático, programación informática y procedimientos de valoración multiparamétrica para comprender y gestionar la opacidad de la información en el mercado inmobiliario italiano, en particular en lo que respecta a la estimación del valor de mercado de las propiedades pertenecientes al sector residencial segmento.

## 1. INTRODUCCIÓN

La transparencia del mercado en un mercado determinado está relacionada con la accesibilidad, la disponibilidad y la calidad y confiabilidad de los datos y la información. Por tanto, la transparencia es uno de los requisitos previos de un mercado perfectamente competitivo y de la condición de equilibrio. Se traduce en la información perfecta para los agentes del mercado (en particular los consumidores) sobre los precios y características de los bienes o servicios.

En un mercado transparente, hay datos de buena calidad disponibles y accesibles tanto para la demanda como para la oferta. En tal caso, los agentes pueden tomar decisiones de compra o producción basadas en los datos disponibles sin requerir un alto

gasto de recursos para su recaudación. Este supuesto es crucial para el buen funcionamiento del mercado y el logro de su eficiencia. Por el contrario, si dichos datos no están disponibles o son muy deficientes, el mercado se define como opaco.

Según Schulte et al. (2005): "Los mercados inmobiliarios pueden describirse como transparentes cuando queda claro cómo funcionan los mecanismos del mercado y las variables detrás de estos mecanismos, es decir, cuando hay tanta información como sea posible disponible en cualquier momento".

John Lang Lasalle (2022) (en adelante JLL) define un mercado inmobiliario transparente como un mercado abierto y claramente organizado que opera dentro de un marco legal y



marco regulatorio caracterizado por un enfoque consistente en la aplicación de reglas y regulaciones.

El Índice Global de Transparencia Inmobiliaria (GRETÍ) 2022 de JLL clasifica a 94 países y 156 ciudades según la transparencia del mercado. Desde 1999, este índice ha considerado varios parámetros, incluida la medición del desempeño de las inversiones, los fundamentos del mercado, la gobernanza de los vehículos cotizados, el entorno regulatorio y legal, los procesos de transacción y la transparencia de la sostenibilidad. Además, las mediciones del desempeño de las inversiones incluyen valoraciones de propiedades. Los mercados altamente transparentes están avanzando gracias a la tecnología, la acción climática, la diversificación de los mercados de capital y los cambios regulatorios. También marca una creciente divergencia entre los principales mercados y otros mercados opacos, con muchos estancados o retrocediendo en el crecimiento de la transparencia.

La sostenibilidad fue el principal impulsor del crecimiento de la transparencia en el índice GRETÍ 2022, y muchos países adoptaron normas obligatorias de eficiencia energética y emisiones para los edificios. Desde la perspectiva de la digitalización, la disponibilidad de nuevos datos específicos y de alta frecuencia está aumentando la transparencia del sector inmobiliario, lo que lleva a una mayor comprensión de cómo funcionan los mercados y los edificios.

Además de los informes de JLL, en los últimos años, las relaciones entre la transparencia y las actividades del mercado inmobiliario, como la inversión, el desarrollo inmobiliario y la valoración, se han examinado en varios artículos (por ejemplo, Farzanegan y Gholipour, 2014; Gholipour y Masron, 2013; Newell, 2016). La bibliografía coincide en que el problema de la transparencia en diferentes países puede dificultar la valoración inmobiliaria, la capacidad de identificar inversiones adecuadas y el diálogo entre los diferentes actores del mercado. Además, una alta transparencia del mercado es la primera línea de defensa contra la incertidumbre. Sin embargo, la transparencia inmobiliaria no es un concepto unívoco.

Este artículo tiene como objetivo discutir cómo la opacidad de la información todavía representa un problema crucial en el mercado inmobiliario italiano, centrándose específicamente en el acceso a la información en la valoración del valor de mercado inmobiliario.

Para una comprensión profunda del nivel de transparencia en el mercado inmobiliario italiano y el posible error que se cometería en las valoraciones inmobiliarias debido a la falta de transparencia, a continuación se propone un modelo para definir y “medir” la opacidad del mercado.

En este sentido, aquí se investiga la transparencia del mercado con la ayuda de técnicas de aprendizaje automático y, por lo tanto, se desarrolla un conjunto de redes neuronales artificiales (RNA) con el objetivo de identificar las principales fuentes de opacidad en el mercado. En particular, las RNA pretenden ser una herramienta de previsión multiparamétrica capaz de estimar el valor de mercado de una propiedad en función de varias características del edificio. Las ANN se entrenan en bases de datos descargadas de sitios web de venta específicos con la ayuda de un procedimiento de descarga automatizado desarrollado ad hoc con el lenguaje informático Python®.

Estos modelos de pronóstico reflejan la “opacidad del mercado” porque se basan en datos sin procesar (información opaca) y los anuncios de venta, como se analizará más adelante, siempre contienen información incompleta, engañosa o incluso incorrecta. En un segundo momento, las bases de datos se vuelven a descargar manualmente. Este procedimiento permite comprobar la exactitud de toda la información reclamada en los anuncios de venta, generando bases de datos más transparentes. Las ANN se implementan en estas bases de datos corregidas, lo que nos permite comprender cuánta opacidad de la información influye en la evaluación del valor de mercado. Además, se aíslan y analizan las fuentes primarias de opacidad.

Como tal, la próxima Sección 2 introducirá y aclarará algunos conceptos que serán útiles para el propósito de este análisis. Luego, la Sección 3 presentará el método adoptado para discutir la opacidad de la información en el mercado inmobiliario italiano, mientras que la Sección 4 presentará los tres estudios de casos prácticos objeto de esta investigación. Sección 5 presentará los procedimientos de capacitación y prueba de las RNA, y la Sección 6 ilustrará el impacto de las variables en el pronóstico del valor de mercado. Finalmente, en la Sección 7 se discutirán las conclusiones de este trabajo.

## 2. FONDO

### 2.1 Transparencia en el mercado inmobiliario

Según la teoría de la transparencia, existe una transparencia estricta en el mercado cuando el intercambio de información se produce con un alto grado de equivalencia entre los participantes del mercado (oferta y demanda), mientras que la falta de transparencia provoca asimetría de información en el mercado (Yun y Chau, 2013). La asimetría de la información se produce cuando algunos participantes del mercado están más informados que otros sobre la información de la transacción (Akerloff, 1970), lo que conduce a resultados distorsionados, mayores costos de transacción y mayor riesgo.

El concepto de transparencia se concibe de manera diferente en el mercado inmobiliario y está vinculado a varios aspectos. Schulte et al. (2005) lo vinculan a la posibilidad de obtener información sobre el mercado inmobiliario y los submercados en los que se divide. Para Linqivitz (2012), la transparencia se refiere principalmente a las transacciones inmobiliarias y cinco aspectos relacionados: información jurídica, fiscalidad financiera, costos de transacción y datos. Otros autores relacionan directamente el nivel de transparencia con el desarrollo tecnológico que impregna el mercado inmobiliario. La transparencia requiere acceso abierto a información nueva y granular con amplia cobertura geográfica pero, al mismo tiempo, implica completa integridad y precisión de las fuentes (Ionas, cu y Anghel, 2020).

En el sector inmobiliario, se cree que un mercado más transparente atraerá más inversiones e inversores (Razali y Adnan, 2012). Más concretamente, cuanto mayor sea el nivel de transparencia inmobiliaria (RET), mayor será el



## Detectando la transparencia de la información en el mercado inmobiliario italiano: un enfoque de aprendizaje automático

Participación y número de inversores extranjeros en bienes raíces (FREI). RET también ha sido investigado en relación con el incumplimiento de las hipotecas – DOM (Gholipour et al., 2020).

Debido a la globalización de las transacciones inmobiliarias y a la presencia de inversores extranjeros, la demanda de (mejor) información sobre datos inmobiliarios ha aumentado significativamente (Farzanegan y Fereidouni, 2014), pero esto no podría aplicarse en todos los mercados, como subraya el último informe de JLL (2022).

Cada mercado inmobiliario muestra su nivel específico de transparencia, y los profesionales en el campo de la valoración inmobiliaria deben hacer frente a la diferente calidad y disponibilidad de las fuentes de datos que pueden utilizar. Por ejemplo, supongamos que un profesional opera en un mercado opaco. En ese caso, debería basar sus valoraciones en datos de mala calidad, lo que daría lugar a valoraciones menos fiables. Si un mismo profesional opera en un mercado transparente, puede disponer de información numerosa y detallada, produciendo una valoración más sólida. La transparencia en el sector inmobiliario debe analizarse teniendo en cuenta las peculiaridades del sector, que determinan el funcionamiento diferente del mercado inmobiliario respecto de cualquier otro mercado (Arnott, 1987).

Además, en el ámbito de la inversión inmobiliaria, la opacidad del mercado causada por la falta de información detallada sobre las características y los precios de los activos conduce a una asignación ineficiente de los recursos y también aumentará el riesgo de inversión. Esto está relacionado con el problema de la asimetría de la información, que involucra tanto a los consumidores, que no pueden conocer todos los aspectos legales, técnicos y económicos del activo que están dispuestos a comprar, como a los inversores, que abordan la inversión con mayor riesgo. Además, la opacidad del mercado hace que las valoraciones de las propiedades sean más inciertas e impide que los compradores ejerzan control sobre la relación precio y calidad de las propiedades que pretenden comprar (Guerrieri, 2011).

Un problema importante relacionado con la opacidad del mercado, la asimetría del mercado y la disponibilidad de información tiene que ver con la valoración inmobiliaria en Italia.

La valoración italiana se ha enfrentado a estándares internacionales desde la introducción de los primeros fondos de inversión inmobiliaria y empresas cotizadas internacionales a finales de los años 1990. Hasta entonces no había sido necesaria una alineación con los demás países, pero la llegada de inversores internacionales requirió un mayor desarrollo en los sectores inmobiliario y de valoración italianos. De hecho, era necesaria una reforma en algunos aspectos, especialmente para codificar el contenido estándar de un informe de valoración y aclarar los pasos para la estimación del valor de la propiedad. Las Normas Internacionales traducidas e introducidas en Italia desde los años 2000 han buscado identificar lenguajes y procedimientos compartidos, mejores prácticas y análisis de información de mercado, contenidos y procesos de una tasación inmobiliaria. Sin embargo, los métodos y enfoques de valoración eran menos preocupantes, ya que ya estaban

utilizado en el sector de tasación italiano. Lo que requirió mayor esfuerzo fue el acceso y el uso de los datos e información en que se basaron las valoraciones. La adopción de normas internacionales era imposible sin establecer primero las reglas para recopilar y analizar datos de mercado.

Sin duda, durante la década de 2000, en relación con el desarrollo de los fondos inmobiliarios y el sólido crecimiento del mercado, se ha producido una marcada mejora en la calidad y cantidad de la información económica.

La construcción de series históricas de precios y el desarrollo de numerosas bases de datos del mercado inmobiliario que proporcionaban precios inmobiliarios en diferentes áreas geográficas han ayudado a difundir buenas prácticas.

En la última década se han divulgado en el mercado inmobiliario una serie de publicaciones, normas y regulaciones sobre la recolección y análisis de datos (UNI/PdR 53:2019, 2019).

La Norma Uni, en particular, tiene como objetivo identificar directrices sobre la metodología de recopilación de datos económicos e inmobiliarios para lograr los objetivos de precisión y exhaustividad de los datos recopilados y lograr transparencia y objetividad en la valoración de bienes inmuebles.

De hecho, es posible observar que el nivel de transparencia del mercado en Italia ha ido mejorando según el Índice Global de Transparencia Inmobiliaria de JLL. Este Índice se expresa como un número de modo que cuanto menor sea el número, menor será la opacidad del mercado. El índice atribuido a Italia en 2001 era de 2,82 (el puesto 23 entre 41 países), mientras que en 2022 es de 2,12. Italia está ahora en el puesto 19.

lugar entre los 94 países analizados por la compañía.

Italia ha perdido dos posiciones desde 2020 y ocupa el puesto 12 entre 29 países europeos.

Las puntuaciones parciales (Fig. 1) para las categorías individuales que componen el Índice cubren el rendimiento de las inversiones, los fundamentos del mercado, los vehículos cotizados, las regulaciones y los aspectos legales, el proceso de transacciones y la sostenibilidad. Las mejores puntuaciones para Italia se refieren a los procesos de transacciones (1,33) relacionados con la información previa a la venta, los procesos de licitación, los estándares profesionales de los agentes y las normas contra el blanqueo de dinero. El peor desempeño se muestra en el índice de sostenibilidad (2,80), que involucra certificaciones de edificios verdes, informes energéticos, estándares de eficiencia y benchmarking energético, informes y estándares de emisiones, arrendamientos verdes, etc. Los índices con el valor más bajo son los que indican mayor transparencia.

Sin embargo, aunque se han realizado mejoras en el campo de la transparencia del mercado en Italia, varios problemas con el intercambio de datos e información todavía afectan y ralentizan los campos de la valoración inmobiliaria y la inversión en este país. Sin duda, el mercado inmobiliario aún necesita continuar el proceso de evolución de la transparencia en sus datos.

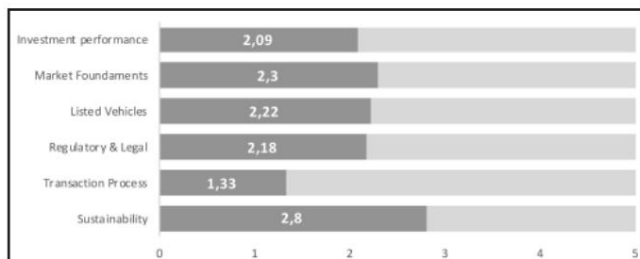


Figura 1 - Desempeño de los subíndices del mercado italiano -  
Fuente: John Lang Lasalle.

## 2.2 Precio versus valor

A los efectos de esta investigación, es decir, para discutir la transparencia de la información en el mercado inmobiliario italiano, es crucial distinguir entre dos conceptos clave que pertenecen al campo de la valoración de propiedades: precio y valor.

Como explica detalladamente (French et al., 2021; Forte e De Rossi, 1974; International Valuation Standards, 2020), el precio es la cifra efectiva a la que se ha vendido una propiedad en un mercado abierto, y es, por tanto, histórico. datos que sólo pueden observarse una vez que se ha producido la transacción. Por el contrario, el valor es una estimación previa de un precio o, en otras palabras, una estimación de la cifra más probable que se pagaría si una propiedad se vendiera en un mercado abierto en la fecha de la valoración. Los precios son hechos históricos, mientras que los valores son estimaciones hipotéticas de precios.

Por lo tanto, los precios y los valores son conceptos profundamente diferentes, que ocurren en diferentes momentos de negociación y que encarnan diferentes roles.

La diferencia entre precios y valores se llama precisión de valoración. Como, nuevamente, se describe bien en (French et al., 2021), la precisión de la valoración se produce cuando la valoración de mercado (la estimación) difiere del precio real alcanzado. La precisión de la valoración representa, por lo tanto, cuán confiable ha sido la evaluación del valor. Esto lleva a la introducción de otro concepto importante, a saber, la variabilidad de la valoración. La diferencia entre dos (o más) valoraciones realizadas por diferentes profesionales sobre el mismo inmueble, al mismo tiempo, en el mismo mercado, se denomina variabilidad de valoración. Por tanto, estamos hablando de precisión de valoración, donde las valoraciones de mercado difieren de los precios de venta reales (ajustados por tiempo). Sin embargo, cuando observamos en qué se diferencia una valoración del mismo activo de otra, esto es variabilidad de valoración.

La precisión de la valoración y la variabilidad de la valoración, hasta cierto punto, indican el "error" que se comete en el proceso de estimación del valor de una propiedad en comparación con el precio de venta real. Una valoración sólo es una estimación de un precio basada en la información disponible cuando se realiza la valoración, no es un cálculo exacto de una cifra. Ninguna fórmula matemática o algoritmo econométrico sofisticado indica un valor preciso de la propiedad. Es un desafío obtener alta

valoraciones fiables de los valores de mercado, ya que las valoraciones siempre se caracterizan por una incertidumbre que la habilidad, la intuición, la competencia y la experiencia del tasador no pueden eliminar.

Las razones pueden residir en varios motivos; algunos ciertamente pueden estar bajo el control del tasador, pero otros no. De hecho, el valor de un activo puede sobreestimarse o subestimarse, por ejemplo, debido a un análisis insuficiente, percepciones personales engañosas, suposiciones previas o una mala interpretación de los datos, lo cual es (en cierto modo) responsabilidad del tasador. Sin embargo, la mayoría de las causas de la baja precisión de la valoración pueden estar fuera del control del profesional.

El artículo ilustra que la primera causa de una baja precisión de la valoración es cuando un profesional opera en un mercado con alta opacidad de información.

Dado que la transparencia/opacidad de la información está relacionada con la disponibilidad y exactitud de los datos en un mercado específico, la precisión y variabilidad de la valoración dependen en gran medida de la confiabilidad de la información proporcionada.

## 2.3 Comparación y comparables

La transparencia del mercado está directamente relacionada tanto con la disponibilidad como con la exactitud de los comparables, en todas las formas, en los que un profesional puede confiar. En la disciplina de valoración italiana, uno de los principios de la valoración inmobiliaria establece que el método es único y se basa en la comparación (Forte y De Rossi, 1974). Incluso en el escenario internacional, la comparación constituye, de hecho, la base de todos los enfoques de valoración reconocidos por los estándares internacionales, es decir, el enfoque de comparación de mercado (o venta), el enfoque de ingresos y el enfoque de costos (Simonotti, 2006).

Los procedimientos de valoración construyen una comparación entre la propiedad que se valora con otros comparables, donde un comparable puede definirse como una propiedad similar al objeto de valoración pero cuyo precio/ingreso/costo es conocido y pertenece al mismo submercado. Además, las características comparables deben ser similares a las del objeto de valoración en términos de características intrínsecas y extrínsecas, tales como, por nombrar algunas, ubicación, nivel de mantenimiento, tamaño, área o tipología de edificio.

Por esta razón, los datos inmobiliarios de los inmuebles utilizados en la comparación son un elemento crucial para el proceso de determinación del valor de mercado, ya sean hechos económicos (precios, costos) o datos técnicos (características cuantitativas y cualitativas). En función de la calidad y cantidad de los datos disponibles, el tasador elegirá el enfoque y el método más apropiado para determinar el valor (French y Gabrielli, 2018).

Dicho esto, varios tipos de fuentes de datos podrían representar un comparable que los profesionales inmobiliarios utilizan para estimar el precio. En relación con el cálculo del valor de mercado de la propiedad y el método aplicado, se

## Detectando la transparencia de la información en el mercado inmobiliario italiano: un enfoque de aprendizaje automático

comparable podría ser una propiedad cuya transacción de venta/alquiler haya tenido lugar recientemente (precio histórico), podría ser una propiedad actualmente en venta (precio de venta), así como cualquier información contenida en datos catastrales, cotizaciones de mercado, estudios de mercado o Tasaciones de inmuebles realizadas por otros profesionales.

Es evidente que diferentes comparables proporcionan diferentes tipos de información, lo que lleva a juicios de evaluación más fuertes o más débiles.

En este sentido, un informe reciente de TEGOVA (francés, 2020) señala cómo los comparables desempeñan un papel crucial en la calidad y fiabilidad de las tasaciones inmobiliarias, y analiza cómo los países europeos están adoptando legislaciones, herramientas y enfoques de valoración opuestos. El informe TEGOVA distingue dos niveles de información de los que los profesionales pueden extraer datos: la información «dura» y la «blanda». El término «información dura» se refiere a datos históricos sobre la venta/alquiler de una propiedad cuando se conocen todos los aspectos económicos (precio) y técnicos (tamaño, estado de mantenimiento, número de habitaciones, pisos, etc.). La recuperación de esta información está relacionada con un documento de transacción (escritura de compraventa), generalmente presentado en notaría, archivos notariales o Registro de la Propiedad.

Por el contrario, la «información blanda» indica fuentes indirectas, informes de empresas de investigación, agencias, bases de datos públicas o revistas. En este caso, el tasador necesita ayuda para comprender cómo se han elaborado o recopilado los datos, y sólo están disponibles algunos detalles generales.

Por lo general, la información blanda también se presenta en forma de datos agregados.

En el sistema italiano de valoración de bienes inmuebles, esta doble clasificación podría identificarse con los términos «fuentes directas» y «fuentes indirectas», incluso si no existe una total adherencia entre estas dos clasificaciones. Las fuentes directas también incluyen los precios de venta contenidos en anuncios y listados de venta. Por el contrario, en el informe de TEGOVA los precios de venta se clasifican como información blanda, junto con las fuentes indirectas.

Esta diferencia entre la clasificación de las fuentes de datos en cada país está relacionada con la disponibilidad de información, la legislación local y el grado de transparencia del mercado. La idoneidad de los comparables no es universal (como muestra el índice de JLL), ni puede usarse de la misma manera incluso dentro de un mismo país (Ionacu et al., 2021), como en el caso de diferentes datos disponibles en ciudades grandes versus pequeñas. Esto implica que las normas internacionales de valoración no deben ser demasiado prescriptivas a la hora de codificar el uso adecuado de comparables, ya que cada fuente de datos puede desempeñar diferentes funciones dependiendo de la transparencia del mercado (Sadayuki et al., 2019) .

### 2.4 Confiar en los precios solicitados

Los vendedores indican el precio de venta con la esperanza de atraer compradores potenciales dispuestos a pagarlo. Este precio se identifica mediante valoraciones, análisis de mercado o inmuebles 1 <https://www.monitorimmobiliare.it/>

agentes. Este precio se coloca al inicio de la venta o negociación de venta y no siempre es un precio definitivo, estando sujeto a un margen de negociación en la mayoría de los casos.

Sólo el precio de venta se coloca al final de la negociación y es una cifra histórica e inmutable.

Los dos precios son, por tanto, diacrónicos.

Los académicos utilizan repetidamente los precios de venta para investigaciones relacionadas con el análisis del mercado inmobiliario y la modelización de precios inmobiliarios (Pozo, 2009; Hayunga y Pace, 2016; Gordon y Winkler, 2016). Además, algunos estudios han relacionado y comparado los precios de venta con los precios de venta en los mercados inmobiliarios (Anglin et al., 2003; Beracha y Seiler, 2014; Knight, 2002). Curto et al. (2015) ilustraron cómo el uso de precios de venta para la valoración de propiedades es típico en Italia debido a la falta de datos sobre los precios de venta. El mismo estudio señala que podría haber problemas al utilizar dichos valores si se utilizaran en la valoración de propiedades. En el verano de 2017 tuvo lugar un animado debate entre académicos y profesionales del mercado inmobiliario en la revista italiana en línea «Monitor Immobiliare»<sup>1</sup>, debatir si se deben utilizar los precios solicitados y cuándo. Las opiniones no son convergentes: muchos profesionales afirman que su uso es necesario cuando los precios solicitados no están disponibles. Otros rechazan el empleo de precios de venta porque no pueden reflejar el valor de mercado de los bienes inmuebles, mientras que otros aceptan una forma de mitigación. Hay mucha distancia entre lo escrito en los manuales, normas y reglamentos, y el mercado inmobiliario, por lo que los operadores deben encontrar formas de adaptación necesarias para su «supervivencia» profesional. En Italia, a pesar de una mayor transparencia del mercado y la construcción de bases de datos históricas, el acceso a los datos sobre compras y ventas sigue siendo limitado, escaso o enormemente caro. Además, existe la idea errónea de que la información histórica de precios sugiere que los inversores no están fijando precios para el futuro. Este no es el caso; un inversor paga una cifra acordada por los rendimientos futuros esperados y, por lo tanto, el precio captura esta visión del futuro. Las valoraciones son indicadores del precio y, por lo tanto, también hacen lo mismo: reflejan las expectativas actuales del mercado sobre el futuro. La valoración es, por tanto, la mejor estimación del precio de venta (futuro) en la fecha de la valoración. Este valor debe considerar el ciclo inmobiliario particular, las tendencias de la demanda y la sostenibilidad de este valor en el período extendido. Además, para ello se requiere información adecuada.

Los profesionales italianos suelen utilizar los precios solicitados como comparables. Los precios de venta están fácilmente disponibles en diferentes mercados inmobiliarios en línea y son gratuitos. Siguiendo las sugerencias de los estándares internacionales, las normas y reglamentos han recomendado qué datos deben utilizarse en las valoraciones de propiedades. Las normas italianas (UNI 11612:2015, 2015) defienden que en



[

En el caso de valoraciones en las que en un período reciente se hayan producido transacciones insuficientes, indetectables y/o poco fiables, los precios de venta sólo podrán tenerse en cuenta de forma residual. Sin embargo, la relevancia de esta información debe quedar claramente definida, analizada críticamente y justificada en el informe de valoración. Esta práctica, sin embargo, va más allá de la excepcionalidad indicada por la Norma.

Sin embargo, es fundamental subrayar que incluso los sitios de venta de inmuebles en línea reflejan la opacidad del mercado italiano. Por ejemplo, es posible comparar los datos disponibles online en Italia con los de los sitios web inmobiliarios estadounidenses. En este último se muestra con qué frecuencia y durante cuántas veces se ha comprado y vendido una propiedad, los diferentes precios de venta, el importe de los impuestos sobre la propiedad y los gastos de funcionamiento. Todo esto es posible gracias a la información de los registros inmobiliarios que, en este caso, está totalmente disponible online.

Por el contrario, los agentes inmobiliarios italianos no pueden acceder en línea a gran escala a las escrituras de compra y, a menudo, la información es escasa. Este también es el caso cuando se consultan los datos de la Agenzia Delle Entrate (Agencia Tributaria). Por ejemplo, en el apartado 'Consulta de precio inmobiliario' está disponible información económica. Aún así, es necesario mejorar la información sobre la unidad inmobiliaria vendida (sólo se puede conocer el tamaño, la categoría catastral y la fecha de la escritura de compra).

Además, incluso si se encuentran transacciones históricas de comparables adecuados, todavía no es posible asegurar la veracidad de la información declarada en una escritura. De hecho, la búsqueda de comparables en los Registros de la Propiedad puede dar lugar a información incompleta o errónea sobre la transacción, el precio de venta o las características del inmueble.

Las características cuantitativas y cualitativas, consideradas variables explicativas de los precios, constituyen un problema de información a la hora de recoger los precios de venta.

Por lo tanto, ¿deberían considerarse los precios de venta como datos inadecuados para las valoraciones inmobiliarias? Lo cierto es que un tasador debe ser consciente del nivel de opacidad que aportan.

En la presente investigación, los autores basan todo el análisis en los precios de venta. En particular, los comparables utilizados para formar a las RNA están representados por los anuncios en línea recopilados en sitios web de venta específicos en Italia.

Sin embargo, antes de debatir si el precio de venta puede ser un valor adecuado para representar el precio de compra (futuro), es necesario comprender cuán veraces son las características descritas en los anuncios y cómo éstas, si se describen incorrectamente por diversas razones, pueden influir en el precio. proceso de estimación del valor de mercado.

¿Cómo pueden las propiedades descritas sumariamente en un anuncio considerarse datos comparables válidos?

¿Pueden las consistencias, el estado de mantenimiento y la

¿Se pueden inferir las características de la propiedad con un grado adecuado de confiabilidad a partir de anuncios comerciales?

¿Qué error cometen los tasadores si utilizan precios de venta en lugar de transacciones históricas?

¿Cuáles son las fuentes de opacidad más importantes en el mercado?

### 3. UN ENFOQUE METODOLÓGICO

La metodología de análisis adoptada en este trabajo de investigación integra procedimientos de programación informática, técnicas de aprendizaje automático y análisis inmobiliario multiparamétrico, con el objetivo de comprender cómo y en qué medida la opacidad en el mercado italiano afecta las previsiones de valores de mercado. Para un mercado determinado (estudio de caso), se definen los siguientes pasos. • En el primer paso, se desarrolla un

software de rastreo automatizado en lenguaje Python®. Un rastreador web es un software que es capaz de navegar y analizar el contenido de las páginas web de forma metódica y automatizada. Para esta investigación se creó un rastreador web con el fin de analizar sitios web de venta italianos específicos y descargar automáticamente los precios de venta de las propiedades inmobiliarias actualmente en venta.

También se descargan varias características del edificio a través del mismo rastreador web junto con los precios de venta.

Este procedimiento permite descargar miles de observaciones fácilmente. Cada observación representa una propiedad en venta cuyo precio de venta y características se conocen. La base de datos, recopilada a través del rastreador web para un mercado específico, generalmente se denomina DBcrawler.

- En el segundo paso se desarrolla una Red Neuronal Artificial basada en el DBcrawler, por lo que se denomina ANNcrawler, con el fin de pronosticar el valor de una premisa en función de sus características en un mercado determinado.

En este sentido, en el campo del aprendizaje automático, las RNA son un sistema computacional capaz de aprender patrones y procedimientos (C' etkovic' et al., 2018). Neural Las redes están hechas de neuronas artificiales y sinapsis artificiales:

las neuronas son las unidades computacionales, mientras que las sinapsis conectan las neuronas entre sí. Las neuronas artificiales se organizan en múltiples capas separadas de modo que las neuronas de entrada se muestran en la capa de entrada y las neuronas de salida están contenidas en la capa de salida. Al mismo tiempo, en el medio, hay múltiples capas ocultas de neuronas. Al igual que en los procedimientos de aprendizaje automático, una red puede "aprender" cómo fluye información de entrada específica desde las neuronas de entrada y genera una salida.

Se puede entrenar una red con cualquier conjunto de datos de información de entrada y salida y, en consecuencia, construir un modelo predictivo (Pittarello et al., 2021). En este artículo, las neuronas de entrada del ANNcrawler son algunas características seleccionadas de la premisa, mientras que la neurona de salida es su valor de mercado correspondiente.

## Detectando la transparencia de la información en el mercado inmobiliario italiano: un enfoque de aprendizaje automático

Específicamente, las RNA desempeñan aquí el papel de una técnica de evaluación del valor de mercado multiparamétrica. De hecho, las RNA se utilizan para analizar varias características de los edificios y pronosticar su valor de mercado. En la función de pronóstico, cada variable independiente (las características del edificio) contribuye de manera diferente a la estimación del precio (la variable dependiente) (Simonotti, 2006). De esta manera, será posible aislar el impacto que cada variable trae al precio y entender qué tipo de error podría producir un nivel diferente de opacidad en la información para cada característica del edificio.

El vector de columna de las neuronas de entrada se denomina rastreador  $[X_r]$ , siendo  $1 \leq r \leq R$  el número total de observaciones en la base de datos. La neurona de salida es un vector de una columna y una fila indicado como rastreador  $[Y_{\text{forecast}}]$ . El rastreador  $[Y_{\text{forecast}}]$ , a través de las capas ocultas, también se puede definir como  $[Y_{\text{forecast}}]_{\text{crawler}} = f([X_r]_{\text{crawler}})$ , es decir, una función del vector  $[X_r]_{\text{crawler}}$ .

Cuando una red se entrena en una base de datos de entrada y salida, la red "aprenderá" cómo se relacionan los datos de entrada con la salida correspondiente. En consecuencia, se evalúa un modelo de previsión. De hecho, durante el entrenamiento en una base de datos, la red cambia iterativamente sus parámetros libres, es decir, los pesos ( $w$ ) y los sesgos ( $b$ ) de la red, hasta que se encuentra el mejor modelo de pronóstico.

En la escala de una sola neurona, durante el entrenamiento, cada  $z$ -ésima neurona recibe una o más entradas numéricas, aquí denominadas  $xz, u$ , con  $1 \leq u \leq U$ , y  $U$  es el número de sinapsis artificiales que ingresan a la neurona. Luego, toda la información recibida se combina dentro del  $z$ -ésimo

neurona usando una función de peso ( $wz, u$ ) y una función de polarización ( $bz$ ), de modo que se produzca una salida numérica ( $Yz$ ). Luego, una función de activación ( $z$ ) transforma el valor de la neurona en un valor de respuesta, como se define mediante la Ecuación 1, en forma de un valor ponderado. suma:

$$\text{neurona } z\text{-ésima}, \quad Yz = z\left(\sum_{u=1}^U w_{z,u} x_{z,u} + b_z\right) \quad (1)$$

Por lo tanto, durante el entrenamiento de la red, los pesos y sesgos se cambian iterativamente hasta que se logra el mejor modelo de pronóstico o, en otras palabras, hasta que el rastreador  $[Y_{\text{forecast}}]$  se vuelve el más cercano al rastreador  $[Y_{\text{expected}}]$ . En esta investigación, el objetivo es el rastreador  $[Y_{\text{expected}}]$ , es decir, el valor de mercado contenido en la base de datos de capacitación (DBcrawler), y el rastreador  $[Y_{\text{forecast}}]$ , como se definió anteriormente, es el valor producido por ANNcrawler. La diferencia entre el valor objetivo menos el pronóstico produce una señal de error denominada  $[err]_{\text{crawler}}$ , como en la Ecuación 2:

$$[err]_{\text{rastreador}} = \text{rastreador } [Y_{\text{expected}}] - \text{rastreador } [Y_{\text{forecast}}] \quad (2)$$

Al entrenar el ANNcrawler previamente definido sobre la información contenida en DBcrawler, el objetivo es minimizar la suma de las señales de error producidas. • En el tercer paso, los Autores pretenden verificar el nivel de transparencia de la información recopilada automáticamente a través del rastreador web. De hecho, además de la inexactitud inherente al uso de precios de venta en las valoraciones de propiedades, es posible verificar que los anuncios de venta también contienen declaraciones falsas, información incorrecta o datos incompletos.

Dado que el artículo tiene como objetivo evaluar cómo y en qué medida esta información engañosa afecta la estimación correcta del valor de mercado inmobiliario, en el tercer paso del desarrollo de la investigación, se recopila manualmente un conjunto de observaciones de control. Por muy largo que sea este procedimiento, representa la única manera de verificar la exactitud de toda la información contenida en los anuncios de venta, de modo que se excluyan los anuncios incompletos cuyos datos no se pudieron verificar, se encuentre información incompleta y se corrijan los datos incorrectos.

Este segundo conjunto de información recopilada manualmente constituye la base de datos de control denominada DBhand. • En el cuarto paso, se prueba el ANNcrawler desarrollado previamente en DBhand. Esta vez, las neuronas

de entrada son las características de las premisas recopiladas y corregidas manualmente, por lo que se denominan  $[X_r]_{\text{mano}}$ , mientras que la neurona de salida es la consiguiente previsión de un valor de mercado  $[Y_{\text{forecast}}]_{\text{mano}}$ . El valor de mercado esperado, o el valor objetivo, es el precio indicado en los anuncios de venta, denominado  $[Y_{\text{expected}}]_{\text{hand}}$ . Claramente,  $[Y_{\text{expected}}]_{\text{mano}} = [Y_{\text{expected}}]_{\text{rastreador}}$ . La diferencia entre el valor de mercado esperado y el pronóstico, como en la Ecuación 3, representa una estimación del error producido debido a las imprecisiones contenidas en los anuncios online (opacidad del mercado):

$$[err]_{\text{mano}} = [Y_{\text{expected}}]_{\text{mano}} / \text{rastreador} - [Y_{\text{forecast}}]_{\text{mano}} \quad (3)$$

## 4. TRES ESTUDIOS DE CASOS PRÁCTICOS EN EL NORTE ITALIA

### 4.1 Descarga de bases de datos: el rastreador web

Para esta investigación, se optó por analizar la opacidad de la información en tres mercados italianos diferentes, específicamente los mercados inmobiliarios en:

- Bolonia, •
- Padua, •
- Trviso.

En opinión de los autores, estas tres ciudades podrían ser buenos estudios de caso representativos porque encarnan tres tamaños de mercado diferentes, donde el tamaño del mercado se refiere aquí tanto al tamaño de la ciudad como al número de elementos

[

que constituyen la oferta y la demanda y el volumen de transacciones. Además, aunque está claro que cada mercado inmobiliario representa un caso específico, ninguna de estas tres ciudades es un caso fuertemente único como podrían serlo, por ejemplo, los mercados de Venecia, Roma o Milán. Por tanto, los resultados pueden ser significativos también para otros mercados similares en Italia.

El primer problema fue comprender cómo recopilar de manera realista los datos y la información necesarios para desarrollar las RNA. Como se presentó en la Sección 3, se desarrolló un software de rastreo automatizado en lenguaje Python para analizar las propiedades inmobiliarias en venta en Bolonia, Padua y Treviso desde sitios web de venta específicos y descargar automáticamente sus precios de venta junto con varias características de los edificios.

La búsqueda en línea del rastreador web debe basarse en la definición de tres dominios de búsqueda web diferentes, cada uno de los cuales es específico de una ciudad. Los tres dominios comprenden propiedades residenciales en venta en Bolonia, Padua y Treviso. Quedan excluidos de este estudio los inmuebles no residenciales, como los comerciales y direccionales, así como los locales en alquiler. En cuanto a la tipología de edificación, se incluyen tanto las construcciones nuevas como las existentes, como áticos, chalets adosados y unifamiliares. apartamentos, chalets pluri/bifamiliares/

En el primer dominio se incluyen en el proceso de descarga las 10 zonas de Bolonia, mientras que en el segundo dominio se incluyen las 14 zonas en las que se divide el municipio de Padua. Finalmente, las 7 zonas de Treviso pertenecen al tercer ámbito.

Para permitir que el rastreador web extraiga la información de cada anuncio de venta, es necesario insertar dentro del código Python el localizador uniforme de recursos (URL) propio de cada anuncio en forma de dirección web, como “https://...”.

Dado que claramente habría sido una operación inviable abrir manualmente todos los anuncios de venta para copiar y pegar sus URL web dentro del código Python, el objetivo era entender cómo producirlos automáticamente.

Se pudo observar que cada anuncio mostraba una URL dada por la combinación de la URL de la página de resultados de búsqueda más un número de serie, donde la página de resultados de búsqueda se puede definir como la lista de todos los anuncios de venta resultantes de la búsqueda en línea en los dominios dados.

Por lo tanto, el rastreador web ha sido programado para extraer primero la URL de la página de resultados de búsqueda, identificar los números de serie del anuncio y crear automáticamente las URL correspondientes.

Después de eso, se implementó la biblioteca de Python “Beautiful Soup” para analizar todas las páginas HTML del anuncio de venta. “Beautiful Soup” es un paquete desarrollado por Leonard Richardson para analizar documentos HTML. Con la ayuda de esta biblioteca, fue posible extraer datos e información de los textos HTML creando un árbol de análisis para todas las páginas analizadas. Se construyó una clase de objetos y funciones en Python para definir el conjunto de información a extraer de cada anuncio.

La clase utilizada en cada dominio se ilustra en la Tabla 1.

Tabla 1 - Clase de objetos y funciones

Elemento de clase o función	uom	Elemento de clase o función	uom	Elemento de clase o función	uom
IDENTIFICACIÓN		PRECIO		Elevar	binario
URL web	texto	Precio	€	Jardín privado	binario
Número de anuncio	número	Superficie del piso	mq	Garaje privado	binario
Código de identificación del anuncio	número	Precio Unitario	€/m²	Jardín compartido	binario
UBICACIÓN		CARACTERÍSTICAS		Plaza de aparcamiento	binario
Zona	texto	norte. baños	número	Sótano	binario
Latitud	coordinar	norte. dormitorios	número	Terraza	binario
Longitud	coordinar	Piso mínimo	número	Automatización de edificios	binario
TIPOLOGÍA DE EDIFICIO		Piso máximo	número	Calefacción central	binario
Departamento	binario	Último piso	binario	Sistema fotovoltaico	binario
Ático	binario	MANTENIMIENTO		Ventilacion mecanica	binario
Casa adosada	binario	Clase energética	número (1 a 10)	Aire acondicionado	binario
Villa plurifamiliar	binario	Nivel de mantenimiento	número (1 a 4)	Fibra óptica	binario
Villa bifamiliar	binario	Año de construcción	número (año)	Chimenea	binario
Villa unifamiliar	binario	CARACTERÍSTICAS		Alarma	binario



## Detectando la transparencia de la información en el mercado inmobiliario italiano: un enfoque de aprendizaje automático

La clase de objetos y funciones definidas anteriormente se ha determinado según un análisis específico de la información disponible contenida en los sitios web de venta de propiedades y según las características más comunes de los edificios utilizados en los procedimientos de valoración de mercado multiparamétricos (Feng y Zhu, 2017; Wang y Xu, 2018). Los atributos incluyen características estructurales/físicas, vecindario y ubicación.

Después de analizar las páginas HTML de los anuncios de venta, se aplicó la biblioteca de análisis de datos "Pandas" en Python. Desarrollada por Wes McKinney, la biblioteca "Pandas" se utiliza para extraer un archivo .xls del procedimiento de rastreo web para mostrar datos e información en forma de tabla. Cada fila de la tabla muestra una observación, mientras que las columnas muestran los elementos de clase descargados por cada anuncio. La base de datos de Bolonia, denominada DB(B)crawler, presenta 2.455 observaciones, en la base de datos descargada de Padua, denominada DB(P)crawler, hay 2.884 observaciones, mientras que la base de datos de Treviso, definida como DB(T)crawler, muestra 1.473 observaciones.

### 4.2 Limpieza de bases de datos: eliminación de registros incompletos y valores atípicos

Después del procedimiento de descarga, fue necesario limpiar las tres bases de datos de observaciones faltantes o engañosas. En primer lugar, se excluyeron de las bases de datos de formación los anuncios incompletos. Luego, también se excluyeron de las bases de datos las observaciones que contenían errores obvios o valores atípicos, por ejemplo, cuando mostraban un precio de venta nulo o una superficie de piso nula.

Específicamente, los porcentajes de observaciones incompletas/engañosas se ilustran en la Tabla 2, divididos por cada ciudad y elemento de clase. Por lo tanto, los porcentajes en la Tabla 2 representan las observaciones que fueron eliminadas del muestreo porque los datos no estaban completos (o correctos) en todas sus clases de objetos. Por ejemplo, la "clase energética" y el "año de construcción" representan el porcentaje más alto de observaciones perdidas debido a la falta de información.

Como tal, las bases de datos de entrenamiento descargadas a través del rastreador web tuvieron que reducirse, respectivamente, a 1.665, 2.122 y 867 observaciones.

## 5. LAS TRES REDES NEURONALES

### 5.1 Entrenamiento con el algoritmo de optimización Cuckoo

Una vez limpiadas las tres bases de datos (DBscrawler) de observaciones inutilizables, fue posible realizar el entrenamiento de los tres ANNscrawler. Los conjuntos de entrenamiento empleados para entrenar las redes se forman seleccionando aleatoriamente el 60% de las observaciones del DBscrawler por cada ciudad. Luego, los conjuntos de selección se definen aleatoriamente

Tabla 2 - Porcentaje de publicidad perdida por clase

Clase	Porcentaje de observaciones perdidas		
	Bolonia	Padua	Treviso
Zona	4,56%	0,43%	5,06%
Latitud	19,85%	2,77%	5,50%
Longitud	1,18%	0,73%	5,57%
Departamento	1,18%	0,00%	0,00%
Ático	1,18%	0,00%	0,00%
Villa plurifamiliar	1,18%	0,00%	0,00%
Villa unifamiliar	1,18%	0,00%	0,00%
Casa adosada	1,18%	0,00%	0,00%
Villa bifamiliar	1,18%	0,00%	0,00%
Precio	1,18%	1,80%	2,65%
Superficie del piso	1,18%	1,32%	0,41%
norte. baños	1,18%	3,16%	3,73%
norte. dormitorios	1,18%	3,47%	3,39%
Piso mínimo	1,18%	2,70%	5,84%
Piso máximo	1,18%	2,70%	5,84%
Último piso	1,18%	0,45%	0,20%
Clase energética	1,18%	22,33%	28,45%
Nivel de mantenimiento	4,56%	6,14%	7,06%
Año de construcción	19,85%	28,54%	21,90%
Elevar	1,18%	0,45%	0,20%
Jardín privado	1,18%	0,45%	0,20%
Jardín compartido	1,18%	0,45%	0,20%
Garaje privado	1,18%	0,45%	0,20%
Plaza de aparcamiento	1,18%	0,45%	0,20%
Sótano	1,18%	0,45%	0,20%
Terraza	1,18%	0,45%	0,20%
Automatización de edificios	1,18%	0,45%	0,20%
Calefacción central	1,18%	0,45%	0,20%
Sistema fotovoltaico	1,18%	0,45%	0,20%
Ventilación mecánica	1,18%	0,45%	0,20%
Aire acondicionado	1,18%	0,45%	0,20%
Fibra óptica	1,18%	0,45%	0,20%
Chimenea	1,18%	0,45%	0,20%
Alarma	1,18%	0,45%	0,20%

[

tomando otro 20% de las instancias restantes por ciudad, mientras que los datos sobrantes crean los respectivos conjuntos de pruebas. Se han considerado tres proporciones de división de conjuntos de datos: 80%-10%-10%, 70%-15%-15% o 60%-20%-20%.

La relación de división definitiva se ha elegido en función del número de muestras y entradas presentes en el conjunto de datos y el modelo. En este caso, las bases de datos son bastante pequeñas. Sin embargo, las numerosas neuronas de entrada hacen que el modelo sea más complejo, por lo que era crucial mantener un número adecuado de observaciones en los conjuntos de selección y prueba.

El procedimiento de capacitación se desarrolló en código Python y se implementó por separado para cada ciudad. Los conjuntos de entrenamiento se utilizan en primer lugar para generar diferentes modelos de ANN; luego, estos modelos se ejecutan en los conjuntos de selección para identificar la ANN que funciona mejor también en el conjunto de selección. Finalmente, el conjunto de pruebas se utiliza para calcular el error en los pronósticos.

Este proceso de entrenamiento se realiza dentro de un procedimiento de optimización que permite probar los diferentes modelos de RNA para minimizar el error en los pronósticos. El

El algoritmo de optimización empleado durante el entrenamiento de las redes es el algoritmo de optimización Cuckoo (Chiroma et al., 2017; Mareli y Twala, 2018). Es un algoritmo de optimización inspirado en la naturaleza que identifica y compara todas las arquitecturas de ANN que muestran un óptimo relativo (es decir, el error mínimo) y elige el óptimo global entre ellas. De hecho, la búsqueda Cuckoo está diseñada para resolver problemas de optimización global, ya que emplea parámetros de conmutación y equilibra paseos aleatorios locales y globales.

En aras de la simplicidad, la red neuronal desarrollada para Bolonia se indicará como rastreador ANN(B), para Padua como rastreador ANN(P) y para Treviso como rastreador ANN(T). Los resultados de las RNA desarrolladas se presentan en la Tabla 3. Los errores medios producidos en el conjunto de pruebas son 9,50% para Bolonia, 11,75% para Padua y 13,55% para Treviso.

Los errores son aproximadamente del mismo orden de magnitud. Sin embargo, podemos ver una correlación inversa con las dimensiones del mercado: cuanto más significativo es el mercado, menor es el error y viceversa.

Tabla 3 - Características de las RNA

NN	N. de entradas	Producción	Número de capas ocultas	Número de neuronas por capa oculta	Funciones de activación	Método de escala	Estrategia de formación
rastreador ANN(B)	32	Valor de mercado	7	184	ReLu	mínimo máximo	significar error al cuadrado
rastreador ANN(P)	32	Valor de mercado	7	144	ReLu	mínimo máximo	significar error al cuadrado
rastreador ANN(T)	32	Valor de mercado	7	144	ReLu	mínimo máximo	significar error al cuadrado

5.2 Prueba de transparencia de la información

Esta sección está dedicada al uso del ANNscrawler desarrollado previamente para discutir cómo la transparencia del mercado (o más bien su opacidad) afecta la confiabilidad de los pronósticos en los mercados italianos.

Para ello, las tres bases de datos se han recopilado manualmente para probar la validez de la información que previamente se había descargado automáticamente a través del rastreador web y se había utilizado para construir las redes. Las bases de datos recopiladas manualmente se denominan DB(B)hand para Bolonia, DB(P)hand para Padua y DB(T)hand para Treviso.

Evidentemente, para identificar el dominio de búsqueda web se han empleado los mismos criterios de selección que para la búsqueda online anterior. Los dominios se limitan nuevamente a propiedades residenciales en venta (excluidos contratos de alquiler u otro tipo de contratos), que comprenden edificios existentes y construcciones nuevas.

Durante esta segunda recopilación de bases de datos, se verificó la exactitud de toda la información y aquellas propiedades cuyos datos no pudieron ser verificados fueron directamente excluidas de DBshand.

Sin duda, esta forma de recopilar datos resultó ser un proceso muy largo y que requirió mucho tiempo. Sin embargo, seguía siendo la única manera de producir una prueba de fuego para comprobar la transparencia del mercado, la exactitud de los datos y la disponibilidad de información.

Las tres bases de datos DBshand se han implementado en ANNscrawler, de modo que las características correctas de las propiedades constituyen ahora las neuronas de entrada para producir una nueva previsión de valor de mercado. Como esperaban los autores, el error producido en las previsiones es mayor que antes. El error medio de Bolonia es del 19,45%, de Padua del 25,57% y de Treviso del 26,97%. Este aumento del error en las previsiones se debe a la opacidad del mercado o, en otras palabras, a la información errónea contenida en los anuncios. Cuanto más grande es el mercado inmobiliario, menor es la tasa de error promedio: los mercados más pequeños muestran menos transparencia en los datos descritos en los anuncios inmobiliarios.

La ubicación resultó ser la principal fuente de opacidad. Sólo unos pocos anuncios comprobados mostraban la ubicación exacta del edificio, mientras que la mayoría estaban colocados en otra calle/distrito equivocado.

Como tendencia general, este problema se acentuó

[

## Detectando la transparencia de la información en el mercado inmobiliario italiano: un enfoque de aprendizaje automático

en Padova y Treviso que en Bolonia. Además, varios anuncios tuvieron que ser incluso excluidos de este análisis porque no proporcionaban suficiente información (fotos y descripciones) para permitir la identificación exacta de la posición de la casa. Sin duda, las razones de esta elevada opacidad residen en la privacidad solicitada por los propietarios y en la estrategia comercial específica adoptada por la agencia inmobiliaria. De hecho, si un comprador potencial no puede encontrar la posición de la casa, será necesaria una intermediación. Sin embargo, la falta de información sobre la ubicación es excesiva y engañosa. Varios anuncios indicaban una ubicación del local, que resultó completamente errónea, confundiendo incluso las zonas central, semicentral y suburbana.

Otro gran foco de opacidad era el nivel de mantenimiento, ya que en ocasiones no coincidía con el resto de datos reportados en los anuncios, como la clase energética, el año de construcción o las instalaciones disponibles. En general, las condiciones de mantenimiento fueron evaluadas con optimismo. En algunos casos, las fotografías mostraban claramente niveles de mantenimiento en condiciones mucho peores que las publicitadas en los anuncios. Sin embargo, esta evaluación es subjetiva y los autores modificaron las condiciones de mantenimiento declaradas sólo cuando eran indudablemente diferentes de las imágenes proporcionadas. Otros errores contenidos a menudo en los anuncios se referían a la identificación y definición de la tipología del edificio, la disponibilidad de un sótano, el nivel del piso, la presencia de un garaje y un jardín privado/compartido. En este sentido, los anuncios online carecían de parte de esta información o no proporcionaban datos consistentes. De hecho, los anuncios contenían tanto una descripción como una tabla, y parte de la información presentada en la descripción era completamente diferente de la de la tabla.

Una vez más, la definición de áticos suele ser vaga y carente de transparencia. La palabra italiana utilizada en el anuncio para indicar ático, es decir, "ático", debería utilizarse para apartamentos de lujo situados en el último piso de un edificio. Sin embargo, en algunos anuncios, el término "ático" también se utilizaba para apartamentos y áticos normales o baratos. Debido a la incertidumbre dada por este parámetro, los autores decidieron señalar sólo el estado del "último piso", sin hacer referencia al nivel de lujo.

Otros errores se referían a las instalaciones: se declaró que los sistemas de aire acondicionado o ventilación mecánica estaban presentes en la casa, pero solo se preparó la estructura general de tuberías y conductos para una hipotética instalación futura.

También hubo otra rama de cuestiones que destacaban la opacidad de la información que no estaba directamente relacionada con las características de los edificios. Los anuncios online sufrieron una alta tasa de modificación/caducidad. Como consecuencia, muchos anuncios fueron eliminados, reintegrados o modificados en tan solo unas pocas semanas. Esto ha llevado a requerir actualizaciones frecuentes y volver a descargar el

bases de datos, y el rastreador web automatizado desarrollado en el marco de esta investigación resultó ser extremadamente útil. Sin embargo, también hubo que realizar múltiples controles manuales constantes, lo que aumentó significativamente la carga de trabajo total.

Otro problema se refería a la distribución asimétrica de la información. Por ejemplo, los anuncios de edificios nuevos solían tener más datos que los de edificios antiguos. Como consecuencia de ello, durante el control manual se tuvieron que descartar más anuncios de edificios antiguos debido a la falta de información esencial. Por esta razón, las tres bases de datos tenían comparativamente más publicidad sobre nuevos edificios.

### 6. DEFINIR LA IMPORTANCIA DE LAS VARIABLES

En el párrafo anterior, se han analizado las principales fuentes de opacidad. Sin embargo, no todos los errores en los anuncios producen el mismo impacto en la predicción del valor de mercado. La opacidad en determinada información es mucho más significativa que en otra. Por esta razón, puede resultar útil determinar qué parámetro de entrada muestra el mayor impacto en el valor de mercado mediante un análisis de importancia de las características.

Entre los enfoques que ayudan a calcular el impacto de las variables en la salida se encuentran los métodos basados en filtros, Wrapper y Embedded (Tatwani y Kumar, 2019).

Los métodos de filtrado se basan en estadísticas univariadas, como el coeficiente de correlación de Pearson, la prueba de chi-cuadrado, la puntuación de Fisher, el umbral de varianza, el índice de dispersión o la diferencia absoluta de medias. Los enfoques basados en Wrapper consideran la selección de un conjunto de características como un problema de búsqueda (Ghosh et al., 2020; Suresh y Narayanan, 2019; Yassi y Moattar, 2014), como la selección de características hacia adelante, la eliminación de características hacia atrás, la Selección exhaustiva de funciones o eliminación recursiva de funciones. Finalmente, los métodos integrados combinan las cualidades de los métodos Filter y Wrapper, como la regularización LASSO y el bosque aleatorio.

La importancia de las características se evalúa aquí utilizando la metodología Random Forest (RF). Se eligió este enfoque porque los métodos integrados son muy precisos y muestran excelentes propiedades de generalización (Siham et al., 2021).

Un Random Forest es un clasificador particular formado por un conjunto de árboles de decisión (clasificadores simples) (Ugolini, 2014), donde un árbol de decisión, en el campo de la informática, es una estructura de datos hecha de nodos y arcos. Un árbol de decisión se lee de arriba a abajo. Los nodos del árbol son los elementos que contienen la información, mientras que los arcos son las conexiones entre los nodos. El nodo inicial es la raíz y no tiene ningún arco de entrada, mientras que los nodos terminales se denominan hojas y no muestra ningún arco saliente.

[

Cada árbol de decisión en un procedimiento de Random Forest se construye (es decir, se entrena) a partir de un subconjunto aleatorio del conjunto de entrenamiento. En este caso, los tres conjuntos de formación son las bases de datos de información de los edificios recopilados respectivamente para Bolonia, Padua y Treviso. Además, cada árbol de decisión también se construye a partir de una extracción aleatoria de las características analizadas (es decir, la información de los edificios). Esta aleatoriedad en la selección de características y observaciones es una parte clave en la construcción de los clasificadores, y está destinada a aumentar su diversidad para disminuir su correlación.

Para definir la importancia de cada característica, es necesario medir cuánto cada característica disminuye su impureza durante el entrenamiento. De hecho, cuanto más disminuye una variable su impureza, más significativa resulta ser esa variable. En clasificación (variables discretas), la impureza viene dada por la impureza de Gini o por la ganancia/reducción de información en entropía. En cambio, en las regresiones (variables continuas), la impureza viene dada por la varianza.

Se definió una matriz de columnas, donde los ejes x representan las características (las variables) y el eje y muestra el objetivo (valor de mercado). Se utilizó la biblioteca "Numpy" para realizar el RF-Regressor y el análisis se realizó en Python. El Regresor RF puede calcular los coeficientes de importancia para cada característica.

El 70% de las observaciones se utilizó como conjunto de entrenamiento, mientras que el 30% restante se utilizó como conjunto de prueba. Durante el procedimiento de RF se construyeron 2.000 árboles y el umbral establecido es 0,75 del valor medio de los coeficientes de importancia calculados. En este caso, la disminución de la impureza de cada rasgo se evalúa como el promedio de las disminuciones dadas por cada árbol que constituye el bosque. De esta forma se estima la importancia final de cada variable. Los coeficientes de importancia calculados por el regresor se muestran en la Tabla 4.

Supongamos que hay información incorrecta en los anuncios sobre los datos más impactantes, como la latitud y la longitud (es decir, la ubicación), el nivel de mantenimiento o la superficie. En ese caso, se cometerá un error considerable en la previsión del valor de mercado. Por el contrario, las variables que menos impactan son las instalaciones y tecnologías como la climatización, la ventilación mecánica, la alarma, el ascensor o la inmótica. Entre las variables de menor impacto se encuentran también el sótano, el jardín compartido y el piso.

7. DISCUSIÓN Y CONCLUSIONES

Este trabajo ha integrado el análisis del mercado inmobiliario con técnicas de evaluación del valor de mercado multiparamétrico, programación informática y procedimientos de aprendizaje automático para investigar la opacidad del mercado en Italia. En primer lugar, un rastreador web automatizado desarrollado en lenguaje Python ayudó a recopilar rápidamente una cantidad considerable de observaciones que describen las propiedades en venta en Bolonia, Padua y Treviso. Sobre la base de estas tres bases de datos, se crearon tres sistemas neuronales artificiales correspondientes.

Tabla 4 - Coeficientes de importancia de RF

Elemento de clase o función	coeficientes RF		
	Bolonia	Padua	Treviso
Variable	%	%	%
Latitud	22,43%	19,69%	15,38%
Longitud	29,30%	19,52%	13,38%
Tipología	0,75%	0,72%	1,05%
Superficie del piso	8,17%	7,35%	7,71%
N. Baños	0,87%	0,93%	1,20%
N. Dormitorios	1,50%	1,12%	0,97%
Piso mínimo	2,03%	2,07%	4,02%
Piso máximo	1,53%	1,63%	1,43%
Último piso	0,61%	0,40%	0,37%
Clase energética	3,19%	3,31%	6,99%
Nivel de mantenimiento	14,76%	20,19%	25,80%
Año de construcción	7,55%	15,46%	10,03%
Elevar	0,62%	1,00%	0,49%
Jardín privado	0,65%	0,35%	0,34%
Jardín compartido	0,50%	0,53%	0,99%
Garaje privado	0,53%	1,01%	4,11%
Plaza de aparcamiento	0,61%	0,59%	0,51%
Sótano	0,81%	0,78%	0,52%
Terraza	0,60%	0,53%	0,78%
Automatización de edificios	0,15%	0,09%	0,12%
Calefacción central	0,62%	0,40%	1,34%
Sistema fotovoltaico	0,08%	0,22%	0,29%
Ventilación mecanica	0,00%	0,16%	0,23%
Aire acondicionado	0,46%	0,29%	0,23%
Fibra óptica	0,59%	0,54%	0,50%
Chimenea	0,50%	0,45%	0,59%
Alarma	0,58%	0,68%	0,64%

Se han entrenado redes en Python para pronosticar el valor de mercado de una propiedad en función de 32 características de entrada, incluyendo, entre otras, ubicación, nivel de mantenimiento, instalaciones y tecnologías, tipología delimitadora, terraza, garaje o jardín. Luego, las tres bases de datos se han vuelto a recopilar manualmente por segunda vez. Este procedimiento permitió controlar cada dato expresado en los anuncios. Se corrigió la información errónea y se excluyeron las observaciones no verificables. Estas tres

[

## Detectando la transparencia de la información en el mercado inmobiliario italiano: un enfoque de aprendizaje automático

Se han implementado bases de datos "limpias" sobre las Redes Neuronales Artificiales desarrolladas anteriormente: el error producido en los pronósticos representa el error en la estimación del valor de mercado debido a la opacidad del mercado (o, en otras palabras, debido a la información incorrecta contenida en los anuncios). Luego, se realizó un análisis de importancia de características basado en la metodología Random Forest.

Como resultado, esta investigación puede ayudar a comprender cómo y en qué medida la opacidad del mercado en Italia afecta la confiabilidad de las valoraciones inmobiliarias. Las redes neuronales artificiales son procedimientos estadísticos de pronóstico adecuados y los modelos de redes neuronales describen con precisión cualquier relación entrada-salida. El uso de un modelo de red neuronal para comparar los resultados de una base de datos opaca con una base de datos "limpia" ha llevado a determinar en qué medida un tasador falla en la mejor estimación debido a la falta de transparencia del mercado.

Además, este análisis multiparamétrico también ha permitido identificar las variables de mayor impacto en la estimación. Esto es crucial ya que las características/información del edificio deben verificarse cuidadosamente para evaluar adecuadamente el valor de mercado. Por tanto, una mayor opacidad en las variables de mayor impacto conducirá a un mayor error en el pronóstico. Por el contrario, una menor transparencia en las variables de menor impacto produciría un error menor en la estimación.

Al final de esta investigación, los autores identifican la necesidad de mejorar el intercambio de datos e información sobre propiedades inmobiliarias en Italia. Los precios históricos de las transacciones deberían estar disponibles, pero también las descripciones de los activos deberían ser mucho más precisas y completas. Además, vender anuncios podría requerir un nivel mínimo de datos antes de considerarse listos para la venta en línea. Además, la información proporcionada en los anuncios debe ser precisa y

correcto, especialmente en lo que respecta a la ubicación de la propiedad. Por último, las ventas y los anuncios deberían ser más transparentes; Incluso diferentes sitios web podrían respetar un diseño mínimo compartido para garantizar la integridad y la claridad.

El propósito es doble: primero, reducir la asimetría de información entre el vendedor y el comprador para que la demanda pueda moverse más conscientemente en el mercado inmobiliario. En segundo lugar, aumentar la transparencia en el mercado, ya que esos datos son utilizados por empresas que analizan el mercado, elaboran informes y publican precios, y por tasadores, que en ocasiones tienen que utilizar precios de venta porque no pueden encontrar precios de transacción.

Además, dado que el mercado inmobiliario se ha vuelto complejo y tiene implicaciones sustanciales para el resto de la economía, todos los operadores deben tener garantías de la calidad de la información recopilada.

El debate aquí analizado gira en torno al concepto de "calidad de la información". Por supuesto, la investigación de mercado no puede prescindir del análisis en profundidad de cada comparable, ya sea de compra, venta o puja. Sin embargo, también parece constructivo centrarse en los procedimientos más que en el tipo de datos: las metodologías apropiadas permiten abordar incluso la información espuria de manera profesional y constructiva, extrayendo ideas significativas.

En el futuro de esta investigación, los autores pretenden aplicar periódicamente la metodología adoptada en otros mercados inmobiliarios italianos para mapear los diferentes niveles de opacidad y comprender si existe una evolución en el tiempo en cuanto al acceso a la información. En particular, será interesante ver si las dinámicas que se han producido en los últimos años (pandemia de Covid-19, la guerra en Ucrania, crisis energética, inflación) pueden, de alguna manera, impactar no solo la dinámica y los precios del real patrimonio, sino también el nivel de transparencia u opacidad.

\* Laura Gabrielli, Departamento de Arquitectura, Universidad IUAV de Venecia correo electrónico: [laura.gabrielli@iuav.it](mailto:laura.gabrielli@iuav.it)

\*\* Aurora Greta Ruggeri, Departamento de Arquitectura, Universidad IUAV de Venecia correo electrónico: [aurora.ruggeri@iuav.it](mailto:aurora.ruggeri@iuav.it)

\*\*\* Massimiliano Scarpa, Departamento de Arquitectura, Universidad IUAV de Venecia correo electrónico: [massimiliano.scarpa@iuav.it](mailto:massimiliano.scarpa@iuav.it)

### Nomenclatura

[err]crawler - señal de error ANNcrawler

[err]hand - señal de error de DBhand

[Xr]crawler - vector de columna de las neuronas de entrada en ANNcrawler

[Xr]hand - vector de columna de las neuronas de entrada de DBmano

[Y\_expected]crawler - valor objetivo ANNcrawler

[Y\_expected]hand - valor objetivo de DBhand

[Y\_forecast]rastreador : vector de una columna y una fila de la neurona de salida ANNcrawler

[Y\_forecast]hand - neurona de salida de DBhand

ANN(B)crawler, - ANN desarrollado en DBcrawler para Bolonia

ANN(P)crawler, - ANN desarrollado en DBcrawler para Padova

ANN(T)crawler - ANN desarrollado en DBcrawler para Treviso



[

ANN/ANN - Red neuronal artificial/Neural artificial  
Redes

ANNcrawler : red neuronal artificial desarrollada sobre la base de DBcrawler

bz – función de sesgo en una neurona

DB(B)crawler : base de datos recopilada a través del rastreador web para  
Bolonia

DB(B)hand - base de datos recopilada manualmente para Bolonia

DB(P)crawler : base de datos recopilada a través del rastreador web para  
Padua

DB(P)hand - base de datos recopilada manualmente para Padua

DB(T)crawler : base de datos recopilada a través del rastreador web para  
Treviso

DB(T)hand - base de datos recopilada manualmente para Treviso

DBcrawler: base de datos recopilada a través del rastreador web

DBhand : base de datos recopilada manualmente

RF - Bosque aleatorio

U - número de sinapsis artificiales que entran en una neurona wz,u

- función de peso en una neurona xz,u -

entradas numéricas en una neurona

Yz - salida numérica en una neurona

z - función de activación en una neurona

## Bibliografía

AKERLOFF GA, El mercado de los "limones": la incertidumbre de la  
calidad y el mecanismo del mercado, *The Quarterly Journal of Economics*,  
84(3), 1970, págs. 488–500.

ANGLIN PM, RUTHERFORD R. AND SPRINGER TM, La compensación  
entre el precio de venta de propiedades residenciales y el tiempo en el  
mercado: el impacto de la fijación de precios, *The Journal of Real Estate  
Finance and Economics*, vol. 26, núm. 1, 2003, págs. 95–111.

ARNOTT R., Teoría Económica y Vivienda. En *Handbook of Regional and  
Urban Economics*, editado por E. Mills, Londres: Elsevier, 1987, págs.  
959–988.

BERACHA E. y SEILER MJ, El efecto de la estrategia de precios de  
cotización en los precios de venta de transacciones, *The Journal of Real  
Estate Finance and Economics*, vol. 49, núm. 2, 2014, págs. 237–255.

C' ETKOVIC' J., LAKIC' S., LAZAREVSKA M., ŽARKOVIC' M., VUJOEVIC'  
S., CVIJOVIC' J. AND GOGIC' M., Evaluación del valor de mercado  
inmobiliario en el mercado europeo mediante la aplicación de redes  
neuronales artificiales, *Complejidad*, vol. 2018, artículo ID 1472957, 2018,  
p. 10.

CHIROMA H., HERAWAN T., FISTER I., FISTER I., ABDULKAREEM S.,  
SHUIB L., HAMZA MF et al., Computación bioinspirada: desarrollo reciente  
sobre las modificaciones del algoritmo de búsqueda del cuco, *Computación  
suave aplicada*, vol. 61, 2017, págs. 149–173.

CURTO R., FREGONARA E. Y SEMERARO P., Comportamiento de  
cotización en el mercado inmobiliario italiano, *Revista Internacional de  
Análisis y Mercados de la Vivienda*, vol. 8, núm. 1, 2015, págs. 97–117.

FARZANEGAN MR Y FEREDOUNI HG, ¿El sector inmobiliario

¿Es importante la transparencia para las inversiones inmobiliarias  
extranjeras?, *Revista Internacional de Gestión Estratégica de Propiedades*,  
vol. 18, núm. 4, 2014, págs. 317–331.

FARZANEGAN MR, GHOLIPOUR HF, ¿Importa la transparencia  
inmobiliaria para las inversiones inmobiliarias extranjeras?  
En t. J. *Estrategia. Gestión de prop.* 18 (4), 2014, págs. 317–331.

FENG J. Y ZHU J., Modelo de regresión no lineal y análisis de opciones  
del precio de bienes raíces, *Dalian Ligong Daxue Xuebao/Revista de la  
Universidad Tecnológica de Dalian*, vol.  
57, núm. 5, 2017, págs. 545–550.

FORTE C. Y DE ROSSI B., *Principi Di Economia Ed Estimo*, Etas., Milán,  
1974.

FRANCÉS N., Precios de mercado. Una investigación sobre el uso de  
pruebas comparables en la valoración de propiedades, *TEGoVA*,  
Asociación del Grupo Europeo de Tasadores, junio de 2020.

FRENCH N., CROSBY N. Y THORNE C., Fijación de precios al mercado:  
valor de mercado: el enigma del malentendido, *Journal of Property  
Investment and Finance*, vol. 39, núm. 5, 2021, págs. 492–499.

FRENCH N., GABRIELLI L., Precios de mercado: valoración de  
propiedades revisada: la jerarquía de enfoques, métodos y modelos de  
valoración, *Journal of Property Investment & Finance*, vol. 36, núm. 4,  
2018, págs. 391–396.

GHOLIPOUR FH, TAJADDINI R., PHAM TNT, Transparencia del mercado  
inmobiliario e incumplimiento de las hipotecas *Investigación en Negocios  
y Finanzas Internacionales* 53 10120, 2020.

GHOLIPOUR FH, MASRON AT, Factores del mercado inmobiliario e  
inversión inmobiliaria extranjera. *J. Economía. Semental.* 40 (4), 2013,  
págs. 448–468.

GHOSH M., GUHA R., SARKAR R. Y ABRAHAM A., Una técnica de  
selección de características de filtro envolvente basada en la optimización  
de colonias de hormigas, *Neural Computing and Applications*, vol. 32,  
núm. 12, 2020, págs. 7839–7857.

GORDON BL Y WINKLER DT, El efecto de los cambios en los precios de  
cotización en el precio de venta de viviendas residenciales unifamiliares,  
*The Journal of Real Estate Finance and Economics*, 2016, págs. 1–31.

GUERRIERI G., L'informazione per l'efficienza e la trasparenza del  
mercato immobiliare: l'esperienza italiana, *Territorio Italia*, n. 1, 2011,  
págs. 88–102.

HAYUNGA DK Y PACE RK, Precios de lista en el mercado inmobiliario  
de EE. UU., *The Journal of Real Estate Finance and Economics*, 2016,  
págs. 1–30.

INTERNATIONAL VALUATION STANDARD COUNCIL IVSC,  
*International Valuation Standards*, Londres, 2020.

IONAS, CU E., ANGHEL I., Mejora de la transparencia inmobiliaria  
mediante la digitalización, *Actas de la Conferencia Internacional sobre  
Excelencia Empresarial* vol.

14(1), julio de 2020, págs. 371–384.

IONAS, CU E., TALTAVULL DE LA PAZ P. Y MIRONIUC M., La  
relación entre precios de la vivienda y transparencia del mercado.  
*Evidencia de los mercados europeos metropolitanos, Vivienda, teoría y  
sociedad*, vol. 38, núm. 1, 2021, págs. 42–71.

JOHN LANG LASALLE, Índice Global de Transparencia Inmobiliaria,

## Detectando la transparencia de la información en el mercado inmobiliario italiano: un enfoque de aprendizaje automático

2022 - Transparencia en una era de incertidumbre, Informe de Transparencia Inmobiliaria, disponible en: [www.joneslanglasalle.com](http://www.joneslanglasalle.com) (consultado en julio de 2022).

LINDQVIST S., El concepto de transparencia en el mercado inmobiliario residencial de la Unión Europea: un marco teórico, *Revista Internacional de Derecho en el Medio Ambiente Construido*, vol. 4, 2012, págs. 99-115.

MARELI M. Y TWALA B., Un algoritmo de búsqueda adaptativo de cuco para optimización, *Computación e informática aplicadas*, vol. 14, núm. 2, 2018, págs. 107-115.

NEWELL G., La cambiante transparencia del mercado inmobiliario en los mercados inmobiliarios europeos. *J. Prop. Invertir. Finanzas*. 34 (4), 2016, págs. 407-420.

PITTARELLO M., SCARPA M., RUGGERI AG, GABRIELLI L. Y SCHIBUOLA L., Redes neuronales artificiales para optimizar proyectos de edificios de energía cero (ZEB) desde las primeras etapas de diseño, *Ciencias Aplicadas*, 11, 2021, p. 5377.

Pozo AG, Una estructura anidada del mercado inmobiliario: evidencia adicional, *Estudios de Vivienda*, vol. 24, núm. 3, 2009, págs. 373-395.

RAZALI MN Y ADNAN YM, Transparencia en las empresas inmobiliarias de Malasia, *Property Management*, 30(5), 2012, págs. 398-415.

SADAYUKI T., HARANO K. Y YAMAZAKI F., Transparencia del mercado e inversión inmobiliaria internacional, *Journal of Property Investment and Finance*, vol. 37, núm. 5, 2019, págs. 503-518.

SCHULTE K.-W., ROTTKE N. Y PITCHKE C., Transparencia en el mercado inmobiliario alemán, *Journal of Property Investment and Finance*, vol. 23, núm. 1, 2005, págs. 90-108.

SIHAM A., SARA S. Y ABDELLAH A., Selección de funciones basada en el aprendizaje automático para la calificación crediticia: una evaluación de filtros y métodos integrados, *Conferencia internacional*

sobre INnovaciones en sistemas y aplicaciones inteligentes (INISTA), 2021, págs. 1-6.

SIMONOTTI M., *Metodi Di Stima Immobiliare*, Flaccovio, Palermo, 2006.

SURESH SMS Y NARAYANAN A., Mejora de la precisión de la clasificación mediante la técnica combinada de selección de funciones de filtro + envoltura, *Conferencia internacional IEEE sobre tecnologías eléctricas, informáticas y de comunicación (ICECCT)*, 2019, págs.

TATWANI S. Y KUMAR E., Comparación paramétrica de varias técnicas de selección de características, *Journal of Advanced Research in Dynamical and Control Systems*, vol. 11, núm. 10 Número especial, 2019, págs. 1180-1190.

UGOLINI M., *Metodologie di apprendimento automatico applicate alla generazione di dati 3d*, 2014, disponible en <https://amslaurea.unibo.it/10415/>.

UNI 11612:2015, Determinación del valor de mercado de las propiedades, 2015.

UNI/PdR 53:2019, Análisis del mercado inmobiliario - Directrices para identificar el segmento de mercado y recopilar datos inmobiliarios, 2019.

WANG A. AND XU Y., Análisis de regresión lineal múltiple del precio inmobiliario, en IEEE (Ed.), *Conferencia Internacional sobre Robots y Sistemas Inteligentes, ICRIS 2018, Changsha (China)*, 2018, págs.

YASSI M. Y MOATTAR MH, Selección de características robusta y estable mediante la integración de métodos de clasificación y técnica de envoltura en la clasificación de datos genéticos, *Comunicaciones de investigación bioquímica y biofísica*, vol. 446, núm. 4, 2014, págs. 850-856.

YUN L. Y CHAU KW, El impacto de la transparencia del mercado inmobiliario en los vínculos entre los bienes raíces indirectos y directos, documento presentado en la Conferencia ERES, Viena, julio de 2013, págs. 3-6.



[www.build.it](http://www.build.it)

**dei**  
TIPOGRAFIA DEL GENIO CIVILE

# Identificar la transferencia informativa en el mercado inmobiliario italiano: una aplicación de aprendizaje automático

Laura Gabrielli\*, Aurora Greta Ruggeri\*\*,  
Massimiliano Scarpa\*\*\*

Chiave de libertad condicional: Analisi del mercato immobiliare;  
Valor de mercado; Precio inicial; Trasparenza;  
Aprendizaje automático; Redes neuronales artificiales

## Abstracto

Esta receta se pone como objetivo para comprender como la transferencia del mercado y la corrección de la información sobre la gripe y los procedimientos de estimación y los procesos de decisión en el mercado inmobiliario italiano. A través del análisis de los tres diferentes mercados inmobiliarios y la verificación de la información relativa a los precios de la oferta, el presente contributo vuole capire se, e in quale dimensione, el impiego de los precios de la oferta en luogo dei reali prezzi di compravendita. Es posible que se cometan errores de valoración y se aumente la incertidumbre en el tiempo y se pregiudicare el proceso de decisión de inversiones negligente.

I risultati della ricerca evidenziano quali sono le fonti primarie di opacità informativa nel mercato immobiliare italiano, classificandolo en base al loro impatto sulla stima del valore immobiliare.

La novedad de esta receta aumenta en el uso integrado de técnicas de aprendizaje automático, programación informática y procedimientos de estimación multiparamétrica para comprender y gestionar la opacidad informativa del mercado inmobiliario italiano, en particular riguardo de la estimación más probable. valore di mercato degli immobili appartenenti al segmento residenziale.

## 1. INTRODUCCIÓN

La misura della trasparenza di un mercato è legata alla disponibilità, completezza, qualità e affidabilità dei dati e delle informazioni. La trasparenza es uno de los presupuestos del mercado perfectamente concordantes y condicionados para su equilibrio y funcionamiento. Cuando el mercado es transparente, los agentes del mercado (en particular, los consumidores) tienen un conocimiento perfecto de los precios y las características de los bienes y servicios.

En un mercado transparente, las informaciones son claras, veritieres, affidabili e complete, ed esse sono accessibili sia da parte dalla domanda e dell'offerta. De tal modo, i soggetti sono in grado di prendere decisioni di consumo o di

Produzione sulla base dei dati a loro disposizione, senza necessitare di un elevato dispendio di risorse per la loro raccolta e analisi. Tale presupposto è cruciale per il corretto funzionamento del mercato e per il raggiungimento della sua Efficienza.

Al contrario, se tali dati non sono disponibili, sono spuri o quantitativamente y qualitativamente scarsi, allora il mercato viene definitivamente opaco.

Secondo Schulte et al. (2005), i mercati immobiliari sono trasparenti quando è chiaro il loro funzionamento e le variabili che influenzano tale meccanismo, ovvero quando, in ogni momento, è disponibile un numero molto ampio di informazioni per gli attori del mercato.

Según John Lang Lasalle (2022) el mercado inmobiliario puede ser considerado transparente cuando es un mercado libre, eficientemente organizado, che opera all'interno di un quadro giuridico e normativo chiaro e coerente.

El Índice Global de Transparencia Inmobiliaria (GRET) 2022 de John Lang Lasalle (de ahora en adelante JLL) ha clasificado 94 países y 156 ciudades en base a un índice de transparencia del mercado. Publicado por la prima volta nel 1999, este indicador toma en consideración diversos parámetros, tra cui la misurazione della performance degli investimenti, i fondamentali del mercato, la Governance dei veicoli quo-tati, il contesto normativo e legale, il Processo di compra - vendita e la trasparenza circa la sostenibilità degli immobili. Tale indice misura, inoltre, la qualità nel mondo delle valutazioni immobiliari. I mercati che nel 2022 si sono affermati per la loro trasparenza hanno incrementato la loro posizione in classifica grazie allo sviluppo tecnologico, alle iniziative per il cambiamento climatico, alla diversificazione dei mercati dei capitali e ai cambiamenti normativi. Además, el índice señala una creciente divergencia tra i mercati più trasparenti e quelli opachi, con muchas di questi ultimi che si fermano o, addirittura, arretrano nello sviluppo della trasparenza dei loro mercati. La principal rueda del crecimiento de la transparencia en los informes GRET del 2022 es el subindicador de sostenibilidad, muchos países han adoptado estándares obligatorios de eficiencia energética y emisiones de los edificios. Desde el punto de vista de la digitalización, la disponibilidad de la base de datos siempre es más affidabili e completi permette la crescita della trasparenza nel settore immobiliare, portando ad una maggiore comprensione delle dinamiche di funzionamento dei mercati y del comportamiento degli edifici.

Da una approfondita analisi della bibliografia esistente, è possibile individuare un filone di ricerca che negli ultimi anni ha esaminato il rapporto tra trasparenza e mercato immobiliare, le valutazioni e gli investimenti, e che è stato poi oggetto di diverse pubblicazioni (Farzanegan e Gholipour, 2014; Gholipour y Masron, 2013; Newell, 2016). Gli studiosi concordano sul fatto che il problema della trasparenza nelle diverse nazioni può rendere complesso il processo di valutazione immobiliare, riducendo la capacità di scegliere gli investimenti más appetibili e compromettendo il dialogo tra i diversi attori del mercato. Solo un'elevata trasparenza informativa consente di gestire l'incertezza e il rischio, e, in general, todos los componentes stocásticas che sono presenti in ogni mercato.

Il presente lavoro si propone di discutere come l'opacità informativa rappresenti ancora un problema cruciale nel mercato immobiliare italiano, focalizzando l'attenzione sulle informazioni impiegate durante el proceso de valoración della stima del valore di mercato degli immobili.

Para comprendere il livello di trasparenza del mercato immobiliare italiano y el error que è possibile commettere nelle stime immobiliari a causa della mancanza di esatta, si propone di seguito un approccio innovativo per la verifica dell'opacità di mercato.

Además, la transferencia viene informada con el uso del nicho tecnológico del aprendizaje automático, y si se activa un conjunto de redes neuronales artificiales (red neuronal artificial, ANN) con la función de individualización de las principales fuentes de opacidad del mercado. En particular, le ANN vengono impiegate come un instrumento de previsión multiparamétrica en grado de estimación del valor del mercado de un inmueble en función delle diverse características descriptivas de ciascun edificio.

La capacitación de ANN si utiliza su base de datos para obtener sitios específicos comerciales de ofertas inmobiliarias disponibles en línea, gracias a un proceso de descarga automático desarrollado ad-hoc con el lenguaje informático Python®. Estos modelos de previsión de la opacidad del mercado se añaden a los datos relativos a los precios de la oferta, como si se ilustraran más adelante, pero la naturaleza contiene información incompleta, debido a errores o adiciones erróneas. En un segundo momento, las bases de datos son estadísticas verificadas puntualmente, para controlar la corrección de todas las informaciones que informan los anuncios negativos de ventas, lo que produce que la base de datos sea más transparente. Le ANN sono poi implementate sui base de datos corretti, consentendo di conoscere quanto le informazioni poco attendibili influenzino the valutazione del valore di mercato.

Además, aislaremos y analizaremos las principales causas de la opacidad de la base de datos.

La ricerca si sviluppa viene segue. Il Paragrafo 2 introduce y discute el concepto di trasparenza e come esso sia trattato nella letteratura. Il Paragrafo 3 illustra l'approccio adottato per individuare l'opacità informativo nel mercato immobiliare italiano, entre el Paragrafo 4 presenta i tre casi-studio su cui verrà applicato il modello. El párrafo 5

Descriva el procedimiento de entrenamiento y prueba del Reti Neurales, mientras que el párrafo 6 chiarisce l'impatto delle variabili del modello sulla stima del valore di mercato. Finalmente, en el párrafo 7 verranno presentamos las conclusiones de este trabajo.

## 2. FONDO

### 2.1 El contrato de transferencia de bienes inmuebles

Secondo i modelli teorici, un mercato è definito trasparente in senso stretto quando todos los agentes del mercado (sia sul versante della domanda che dell'offerta) possono donare lo stesso grado di informazioni y la stessa possibilità di accesso alle fonti. Di converso, la mancanza di trasparenza causa asimmetria informativa nel mercato (Yun e Chau, 2013). L'asimmetria informativa si verifica quando alcuni attori sono più informati di otros durante tutte le fasi della compravendita immobiliare (Akerloff, 1970), e ciò determina distorsioni informative, maggiori costi di transazione e un incremento della rischiosità degli investimenti.

Sul significato di "trasparenza" non vi è convergenza, soprattutto se legato al particolare contesto immobiliare.

Schulte et al. (2005) associano la trasparenza alla possibilità di raccogliere informazioni sul mercato immobiliare e



## Identificar la transferencia informativa en el mercado inmobiliario italiano: una aplicación de aprendizaje automático

sui sotto-mercati in cui esso può essere segmentato per tipologia, funzioni, utenza, ecc. Según Linquivitz (2012), la trasparenza concerne principalmente le transazioni immobiliari e gli aspetti legali, la tassazione, i costi e le informazioni ad esse legati. Otros autores collegano directamente el nivel de transparencia al desarrollo tecnológico que pervade el mercado inmobiliario: la transparencia requiere un libre acceso a informaciones puntuales con amplia cobertura geográfica, pero al mismo tiempo implica completitud y precisión de las fuentes informativas (Ionas, Cucu e Anghel, 2020).

Nel settore immobiliare, si ritiene che un mercato più trasparente attragga più investimenti e investitori (Razali e Adnan, 2012). En particular, maggiore è il livello di trasparenza (Transparencia Inmobiliaria - RET), maggiore è la partecipazione di investitori stranieri attratti dal settore immobiliare (Inversores extranjeros en bienes raíces FREI). La RET è stata studiata in relazione anche all'insolvenza dei mutui (Gholipour et al., 2020).

A causa della globalizzazione delle transazioni immobiliari e della presenza di investitori stranieri, la richiesta di (migliori) informazioni sui dati immobiliari è aumentata in modo significativo (Farzanegan e Fereidouni, 2014), ma ciò non può essere applicato indistintamente in tutti i mercati, subsiguiente al último informe de JLL del 2022.

Infatti, ogni mercato immobiliare presenta un livello di trasparenza specifico e i professionisti che operano nel campo della valutazione immobiliare devono confrontarsi con la diversa qualità e disponibilità delle fonti di dati che possono utilizzare. Si una opera profesional en un mercado opaco, debe basarse necesariamente en los datos de la estimación de la estimación, perjudicando la declaración de la estimación de la estimación. Se lo stesso professionista operasse in un mercato trasparente, con la possibilità di disporre di numerose informazioni puntuali, la estimación del más probable valor del mercado potrebbe risultare más robusta ed attendibile. Quindi il concetto di trasparenza nel settore immobiliare deve essere analizzato in relazione alle peculiarità del segmento specifico, che determinano il diverso funzionamento del mercato immobiliare rispetto a qualsiasi altro mercato (Arnott, 1987).

Además, en el campo de las inversiones inmobiliarias, la opacidad del mercado causada por la falta o escasez de informaciones detalladas sobre las características y los precios de los activos causada una ineficiente asignación de los recursos y un aumento de la riesgo de la inversión. Ciò è legato al problema dell'asimmetria informativa, che coinvolge sia i consumatori, che non possono conoscere tutti gli aspetti legali, tecnici ed economici del bene che intendono acquistare, sia gli investitori, che affrontano maggiori rischi durante l'investimento. Además, la opacidad del mercado hace más inciertas las valoraciones de los inmuebles ostacolando porque la verificación de la relación entre la calidad y el precio de los inmuebles de la parte de los adquirentes (Guerrieri, 2011).

Un problema rilevante relativo a la opacidad del mercado,

all'asimmetria del mercato e alla disponibilità di informazioni riguarda la valutazione immobiliare in Italia.

L'estimo in Italia si è allineato con il contesto internazionale fin dall'introduzione dei primi fondi di investimento immobiliare e di società quotate multinazionali verso la fine degli anni '90. L'arrivo e l'interesse degli investitori internazionali ha imposto un confronto tra le diverse impostazioni teoriche della dottrina estimativa e delle metodologie valutative. Gli Standard

Internazionali tradotti e diffusi in Italia a partire dagli anni 2000 hanno introdotto standard etici e professionali, buone pratiche, principi, regole e procedimento per garantire stime oggettive, ripercorribili e trasparenti. Se la disciplina estimativa italiana ha una fisionomia ben delimitata, il cui impianto teorico poggia su una serie di postulati aventi validità generale, gli Standard Internazionali hanno invece cercato di indivi-

Due linguaggi e procedimento condivise, regole di natura metodologica e applicative, mejores prácticas per la rilevazione delle informazioni del mercato, dei contenuti e i passaggi di una valutazione immobiliare.

Senza embargo, la adozione di Standard Internazionali è molto completa, se non è impossibile, senza prima stabilire delle regole per la raccolta e l'analisi delle informazioni di mercato. Ciò ha richiesto uno sforzo maggiore in relazione all'accesso, all'utilizzo e alla diffusione dei dati e delle informazioni immobiliari.

Senza dubbio, desde el año 2000, en concomitanza con lo sviluppo dei fondi immobiliari italiani e la rapida crescita del mercato, si è assistito a un netto miglioramento della qualità e della quantità di informazioni economiche.

Grazie alla creazione di serie storiche dei prezzi, alla divulgazione di studi e di ricerche, alla crescita della copertura territoriale degli osservatori del mercato in grado di fornire quotazioni per diversa categoria, tipologia e microzona si sono diffuse una serie di Buena práctica para cuanto concierne al análisis del mercado.

En particular, en el último decenio, la Normativa UNI (UNI/PdR 53:2019, 2019), insieme ad altre pubblicazioni e regolamenti, ha focalizzato l'attenzione sulla raccolta dei dati immobiliari e la loro analisi. La Norma Uni, in particular, ha venido obbiettivo l'individuazione di linee guida sulla metodologia di rilevazione dei dati economici e immobiliari, per conseguire obiettivi di precisione, completezza dei dati rilevati, e conseguire trasparenza e oggettività nella stima immobiliare.

Como consecuencia de este nuevo concurso de mercado, es posible notar esto, segundo el Índice Global de Transparencia Inmobiliaria de JLL, el vivo de transparencia del mercado en Italia cada vez más años atrás. El índice de JLL viene expresado a través de un número de índice: más bajo y cuanto va-loro, menos opaco y el mercado. L'indice attribuito al mercato italiano nel 2001 era pari a 2,82 (colocado al 23° posto su 41 Paesi, e considerato un mercato semi - trasparente).

El índice de cuentas de 2022 es del período 2,12 (mercado transparente), llevando Italia al puesto 19° de 94 Naciones ahora incluye nel paniere monitorato dal JLL. L'Italia ha perso due posti rispetto al 2020 (quando si collocava al 17° posto), ed è ora

[

al 12° posto tra i 29 Paesi Europei analizzati.

I punteggi parziali (Fig. 1) per le singole categorie di cui si componen l'Indice di JLL riguardano la Performance dell'investimento, i Fondamenti del mercato, i Veicoli di investimento quotati, le Leggi e Regolamenti, il Processo di transazione immobiliare e la Sostenibilità. I punteggi migliori raggiunti dall'Italia riguardano il Processo di compravendita (1,33), in particolar modo le informazioni pre-vendita, la commercializzazione del bene nel mercato, gli standard professionali degli agenti immobiliari e la normativa antiriciclaggio. La performance peggiore si registra nell'indice di sostenibilità (2,80), che riguarda le certificazioni di bioedilizia, la rendicontazione energetica, gli standard di efficienza, il calcolo delle emissioni, i contratti di locazione verdi, etc. En los gráficos, los indicadores con el valor más bajo son los que indican una mayor transferencia en el mercado italiano.

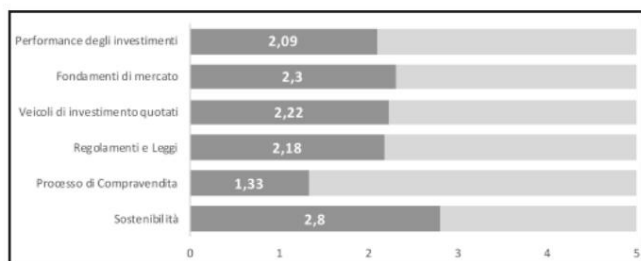


Figura 1 - Andamento dei sottoindici del mercato italiano -  
Fonte: John Lang Lasalle.

Tuttavia, anche se in Italia sono stati fatti grandi passi in avanti a partire dagli anni '90 ad oggi, la difficoltà di ripetición y difusión de la información y los precios del mercado impattano sulla qualità delle valutazioni immobiliari, sulle analisi di mercato e sulle valutazioni delle opportunità di investimento nel nostro Paese. Senza dubbio, il mercato immobiliare deve continuare il processo di crescita del livello di trasparenza cominciato oramai da qualche decennio.

## 2.2 Precio de compraventa y valor de mercado

Ai fini della presente ricerca, che ha venido obietivo l'analisi e la comprensione della trasparenza dei dati e dell'uso delle informazioni nel mercato immobiliare italiano, appare fondamentale distinguere tra due concetti chiave nell'ambito delle valutazioni immobiliari: prezzo e valore.

La differenza tra prezzo e valore è puntualmente illustrata da (Forte e De Rossi, 1974; French et al., 2021; International Valuation Standard Council, 2020): il prezzo rappresenta la quantità di denaro realmente pagata da un acquirente per un immobile che viene scambiato nel mercato, un dato storico che può essere osservato solo a transazione avvenuta. Il valore di mercato di un bene è, invece, un giudizio di stima, una previsione, l'importo al quale l'immobile potrebbe essere venduto alle normali condizioni di mercato alla data della stima. Il prezzo, pertanto, è un dato storico definito dal mercato, una quantità di denaro effet-

tivamente pagata per l'acquisto di un immobile, mentre il valore è un giudizio di stima che esprime, in termini ipotetici, il futuro prezzo del bene. Se ne deduce che prezzo e valore sono concetti profondamente differenti, che si verificano in tempi differenti della negoziazione e assumono ruoli diversi nel mercato.

La diferencia monetaria entre precios y valores es medida por la precisión de la estimación. Como se describe en (French et al., 2021), se define la precisión de la valoración como la medida en la que la estimación del valor de mercado difiere del precio de compraventa alcanzado por la negociación, misurando, en algún modo, la fiabilidad de la valoración inmobiliaria. Por el contrario, la diferencia entre los tiempos de los inmobiliarios de medesimo beneficio, redatada por profesionales diversos en el mejor momento y en las mejores condiciones del mercado, es claramente variable. Si habla, por tanto, de la precisión de la estimación cuando los valores de mercado identificados difieren de los precios de compraventa efectivos, mientras se habla de variabilidad de las estimaciones cuando se comparan valoraciones del mismo bien efectuadas por sujetos diversos. La precisión y la variabilidad de la valoración, en una cierta medida, indican ambas el «error» cometido en el proceso de estimación de la magnitud económica respecto al precio de venta efectivo.

La estimación del mayor valor probable del mercado es una indicación de un posible precio futuro basado en la información disponible en el momento de la estimación. Esto, sin embargo, no se reduce a una mera elaboración matemática de los datos a través de un más o menos sofisticado algoritmo, que lleva a la formulación exacta de la entidad económica (precio, o costo). Aunque hoy en día se utilizan modelos econométricos capaces de reproducir con suficiente asistencia y mecánica del mercado, ninguno de ellos podrá llegar a la indicación de un precio «cierto», en tanto, por su naturaleza, la estimación tiene un carácter probabilístico, no determinístico, y las estimaciones son siempre connotadas por un cierto nivel de aleatoriedad, que habilidad, intuición, competencia y experiencia del valorador no pueden eliminar completamente.

Las razones de los errores pueden ser diversas, algunas controlables por el evaluador, otras no. En efecto, el valor del bien puede ser sobrestimado o subestimado debido a una escasa información de mercado, a convicciones personales erróneas o erróneas, a hipótesis equivocadas, o incluso a una interpretación errónea de los datos encontrados en el mercado. Sin embargo, al menos algunas causas que pueden ser directamente responsables de la valoración, la escasa corrección de las estimaciones podría estar relacionada con factores no controlables, como el riesgo, la incertidumbre y la acción de los componentes estocásticos presentes en el mercado.

Como se ilustra en el presente informe, la causa principal de una baja corrección de la valoración es que un profesional debe operar en un mercado con un nivel elevado de opacidad informativa. Dado que la transparencia/opacidad informativa es ligada a la disponibilidad y a la corrección de los datos obtenidos en un mercado específico, la precisión y la variabilidad de la valoración dependen fuertemente de la fiabilidad de la información proporcionada.

[

## Identificar la transferencia informativa en el mercado inmobiliario italiano: una aplicación de aprendizaje automático

### 2.3 Comparación de datos de mercado

La trasparenza del mercato è direttamente correlata alla disponibilità, alla correttezza e all'affidabilità dei campioni significativi dei prezzi di vendita e delle caratteristiche dei beni analoghi all'oggetto di stima che il professionista può impiegare durante la perizia. Nella disciplina estimativa italiana, uno dei principi del estimado afferma che il metodo di stima è unico ed è basato sulla comparazione (Forte y De Rossi, 1974).

Además del concurso internacional y de todos los internados de la escuela de matrices anglosajona, la comparación constituye el fundamento de todos los procedimientos de estimación de ciencias ricas de los estándares internacionales, sobre el enfoque de comparación de mercados, el enfoque de ingresos y el enfoque de costos (Simonotti, 2006).

La comparación si se basa en el enfrentamiento de los inmuebles de la estimación con bienes similares, por los que resultan notables el precio, el ingreso o los costos, recaídos en el mismo segmento de mercado. Oltre alle grandezze economiche, debbono essere note anche le caratteristiche estrinseche ed intrinseche degli immobili che compongono il campione significativo di comparazione, quali, per citarne alcune, l'ubicazione, il livello di manutenzione, la superficie, l'affaccio, il livello di piano o la tipologia edilizia.

Según la historia, los datos inmobiliarios de los elementos empleados en la comparación constituyen un elemento crucial para la estimación, ya sean económicos (precios, costos) o datos técnicos (características cuantitativas y cualitativas). In base alla qualità e alla quantità dei dati disponibili, il valutatore sceglierà il procedimento (e, successivamente, il modello matematico) più idoneo per giungere alla formulazione del giudizio di stima (French e Gabrielli, 2018).

Diversas tipologías de información (precios efectivos, precios solicitados, precios de locación, costos) pueden representar datos comparables que los profesionales del sector inmobiliario utilizan para formular un juicio de estimación. In relazione allo scopo della stima e al procedimento adottato, i dati relativi ad un campione significativo possono essere un immobile la cui transazione di vendita/affitto è avvenuta di recente (prezzo storico), un immobile attualmente in vendita (prezzo richiesto), così come qualsiasi informazione contenuta nei dati catastali, nelle quotazioni di mercato, nelle indagini di mercato o nelle perizie effettuate da altri professionisti. È chiaro che i diversi dati immobiliari disponibili, ovvero i prezzi o redditi di mercato, possono dare un diverso livello di qualità informativo, da cui dipendono (e derivano) giudizi di stima più o meno robusti.

A este respecto, un reciente informe de TEGOVA (francés, 2020) subraya como comparables juegan un papel crucial en la calidad y en la fiabilidad de las valoraciones inmobiliarias. Lo escrito, inoltre, illustra come diversi Paesi europei stiano adottando legislazioni, strumenti e approcci molto diversi relativamente alle valutazioni immobiliari. Il rapporto TEGOVA distingue due livelli di informazioni da cui i professionisti possono desumere i dati per la

stima, ovvero informazioni «duro» y «suave». Il termine «hard information» si riferisce ai dati storici relativi alla vendita e locazione di un immobile quando sono noti tutti gli aspetti economici (prezzo) e tecnici (dimensioni, stato di manutenzione, numero di stanze, piano, ecc.). Il reperimento di queste informazioni è legato a un documento di transazione immobiliare (atto di compravendita o locazione), solitamente depositato presso gli studi notarili, archivi notarili, la Conservatoria o l'Agenzia delle Entrate.

Por el contrario, las «informaciones blandas» indican los informes de la sociedad de arrendamiento, los datos de las agencias, los bancos de datos públicos o los especializados. En este caso, el evaluador no puede conocer los datos de si han sido recolectados y elaborados, pero son disponibles pocas indicaciones generales alrededor de la determinación del valor. Di solito, le informazioni soft prendono forma di dati di mercato aggregati, pur a differente livello (locale, regionale, nazionale).

Nella tradizione estimativa italiana, esta doble clasificación podría ser identificada a través de la articulación en «fuentes directas» e «fuentes indirectas», anche se non c'è una totale convergenza tra le due classificazioni. Las fuentes directas comprenden también los precios de los inmuebles que se ofrecen en el mercado y los contenidos en los anuncios de venta y en las inserciones online. Por el contrario, en el informe TE-GOVA, los precios de venta son clasificados como información suave, insieme alle informazioni provenienti da rapporti e ricerche di società che realizzano analisi di mercato.

Esta diferencia entre la clasificación de las fuentes informativas en cada País es fuertemente condicionada por la disponibilidad de datos puntuales, por la legislación local y por el grado de transparencia del mercado inmobiliario.

Sin embargo, la calidad de los datos inmobiliarios no es análoga a cada país (como ampliamente demostrada en los índices de JLL), pero no lo es para todos los internos de cada nación, como en el caso de las diversas informaciones disponibles en las grandes ciudades, respecto a aquellas de pequeñas dimensiones (Ionascu et al., 2021). Ciò implica che gli Standard di Valutazione Internazionali non dovrebbero essere eccessivamente prescrittivi nel codificare l'uso appropriato dei dati comparabili, poiché ogni fonte informativa può svolgere ruoli diversi a seconda del livello di trasparenza del mercato locale (Sadayuki et al., 2019).

### 2.4 La stima attraverso i prezzi richiesti (precio de venta)

I prezzi richiesti o di prezzi offerti<sup>1</sup> (en inglés, pidiendo precios) sono i prezzi indicati dai venditori nella speranza di

<sup>1</sup> In italiano i termini «prezzi richiesti» o «di offerta» vengono impiegati in modo interscambiabile. En español gli «precios de venta» (o incluso, más raramente, precios de lista) sono i prezzi richiesti dal venditore, mentre gli «offerprices» sono i prezzi proposti dall'acquirente al momento della trattativa e che riflettono la sua disponibilità a pagare per un dato bene.

[

attirare potenziali acquirenti disposti a pagare tale importo. El precio viene individualizado a través del tiempo o análisis del mercado o, ancora, dagli stessi agenti immobiliari. Il prezzo di offerta è collocato all'inizio della vendita o della trattativa di vendita e non sempre è un prezzo definitivo, essendo soggetto, nella maggior parte dei casi, ad un margen di trattativa. Si tratta, quindi di un prezzo temporaneo e unilaterale. Solo el precio de compravendita, invece, si colloca alla fine della trattativa tra le due parti, acquirente e venditore, ed è un dato storico e inmodificabile. I due prezzi sono, pertanto, diacronici.

Los precios de las ofertas son impiegati in diversos ricerche per analizzare i trend e le dinamiche del mercato immobiliare, e per la costruzione di indici (Pozo, 2009; Hayunga e Pace, 2016; Gordon e Winkler, 2016). Además, alcuni studi hanno messo in relazione e confronto i prezzi di richiesta con i prezzi di vendita nei mercati immobiliari (Anglin et al., 2003; Beracha e Seiler, 2014; Knight, 2002). El artículo de Curto et al. (2015) ilustra el uso de los precios de la oferta para la valoración de los inmuebles sia muy difuso en Italia, e tale diffusione sia provocata principalmente dalla scarsa accessibilità ai dati relativi ai prezzi di compravendita. Lo stesso studio rileva che l'utilizzo di tali valori può rendere más difícil la valoración de los inmuebles.

Nell'estate del 2017 si è svolto un animado dibattito tra studiosi e professionisti del mercato immobiliare sulla rivista online italiana «Monitor Immobiliare». Oggetto della discussione sono stati proprio gli pedir precio e il loro possibile uso nelle valutazioni. Le opinioni non sono convergenti: alcuni professionisti affermano che l'uso si rende necessario quando i prezzi richiesti non sono disponibili. Altri rifiutano el uso del precio de la oferta perché non può in alcun modo riflettere il valore di mercato dell'immobile, entre otros ancora propongono una especie de compromiso messo tra le due posizioni opposte. En general, nella discussione viene sollevato il problema che tra ciò che è scritto nei manuali, nelle norme e nei regolamenti, e il mercato immobiliare vi sia molta distancia, per cui gli operatori devono trovare forme di adattamento necessarie alla loro “sopravvivenza” profesional. Infatti, nonostante la maggiore trasparenza del mercato e la costruzione di banche dati storiche, in Italia l'accesso ai dati sulle compravendite è tuttora limitato o estremamente costoso. Non mancano i difensori dell'impiego degli pidiendo precio en el tiempo, aducendo como motivo il fatto che i prezzi di compravendita, essendo storici, riflettono solo il pasado e precludano lo sguardo verso il futuro. Tuttavia, un acquirente, un investitore, è disposto a pagare una somma di denaro al momento della compravendita in grado di riflettere le attese Future di valore e rendimento. Quindi i prezzi di compravendita sono in grado di riflettere le attuali aspettative rispetto all'evoluzione del mercato nel futuro. La valoración es la mejor estimación del (futuro) precio de compra de todos los datos de la estimación, y el valor debe tenerse en cuenta del ciclo inmobiliario particular, el andamento del hogar y la sostenibilidad del valor en el medio período. E per fare esto es necesario un adecuado patrimonio informativo.

È pratica diffusa dai valutatori e professionisti italiani el uso dei prezzi di offerta individuati nei siti immobiliari specializzati. L'ampia reperibilità, il facile accesso e la gratuità sono tra le cause del long impiego dei prezzi di immobili offerti in vendita. Gli Standard Internazionali, insieme a diversi regolamenti e norme, hanno indicato quali dati potrebbero essere impiegati nelle stime. En Italia, la Norma UNI del 2015 (UNI 11612:2015, 2015) ha ammesso, pur in alcuni casi residuali, l'impiego degli precio de venta. En particular, en el caso de valoraciones de inmuebles en la zona en que no siano avvenute suficientes transazioni, e non sia possibile individuare prezzi di compravendita recientes di un campione significativo analogo al bene da stimare, o che tali prezzi di compravendita individuati risultino inaffidabili, potranno essere impiegate le richieste di prezzi di immobili simili offerti in vendita. L'impiego e la rilevanza di tali informazioni deve essere chiaramente definita, analizzata criticamente e giustificata nel rapporto di valutazione fornito al Committente. Tale pratica, però, si stende oltre l'eccezionalità indicata dalla Norma.

Sin embargo, è importante sottolineare che anche i siti di vendita immobiliare disponibili online riflettono l'opacità connessa al mercato italiano. A titolo di esempio, è possibile confrontare i dati disponibili online in Italia con quelli presenti nei siti, per esempio, di vendita immobiliare statunitensi. Da esto ultimi è possibile reperire informazioni di quali y quante volte un immobile è stato compravenduto, i diversi prezzi di compravendita, l'ammontare di tase di proprietà o una stima dei costi di gestione. La disponibilidad de datos es consecuencia de un sistema transparente que permite todas las informaciones con tenute nel catasto, che, en este caso, sono totalmente disponible online a chiunque.

Di control, gli atti di compravendita in Italia non sono fruibili telematicamente su larga scala dagli operatori del settore immobiliare e spesso le informazioni contenute in essi sono scarse. Anche interrogando i dati dell'Agenzia delle Entrate, nella sezione “Consultazione valori immobiliari dichiarati”, è possibile disporre delle informazioni economiche (prezzo), ma i dati circa l'unità immobiliare compravenduta sono scarse (si hanno indicazioni limitatamente a dimensione, categoria catastale, e data dell'atto di compravendita).

Además, si es posible registrar una información histórica relativa a las analogías de las estimaciones, no es posible garantizar la veracidad de la información de las transacciones negligentemente compradas. La ricerca di campioni significativi di immobili compravenduti presso la Conservatoria può portare a informazioni incomplete o, peggio, errate relative alla compravendita, al prezzo o alle caratteristiche dell'immobile. Le

caratteristiche quantitative y cualitativa quindi, consideradas como variables esplicative del precio, sono un problema informativo nella raccolta dei prezzi reali di compravendita.

¿Gli pedir precio devono quindi essere considerati dati non adeguati per le stime immobiliari? Quello che è

[



## Identificar la transferencia informativa en el mercado inmobiliario italiano: una aplicación de aprendizaje automático

certo è che un valutatore deve essere consapevole del livello di opacità che essi comportano.

Obiettivo di esta ricerca è l'analisi e la verifica della correttezza degli pidiendo precio qualora fossero impiegati como única fuente en la estimación del más probable valor de mercado. En particular, los inmuebles comparabili y utilizados per il Training delle Reti Neurali sono rappresentati dagli annunci online raccolti suspecifici siti web specializzati nella vendita di immobili in Italia.

Comunque, prima di esaminare se il prezzo richiesto possa essere un valore congruo a rappresentare il (futuro) prezzo di compravendita, è necesario capire quanto siano veritiere le caratteristiche descritte negli annunci pubblicitari, e come queste, se descritte erroneamente por diversos ragioni, possano influenzare il procedimento di stima del più probabile valore di mercato.

In che modo gli immobili descritti sommariamente in un annuncio immobiliare possono essere considerati dati attendibili per le stime?

¿Le consistente, lo stato di manutenzione e le caratteristiche dell'immobile possono essere dedotte con un adeguato grado di affidabilità dagli annunci immobiliari?

¿Cuáles son los errores que se cometen en los tasadores que se utilizan en los precios más ricos de los vendedores en el lugar de los datos de las transacciones históricas?

¿Cuáles son las principales fuentes de opacidad del mercado?

### 3. ONU APROCCIO METODOLOGICO

La metodologia di analisi adottata in esta ricerca integra procedimento di programmazione informatica, tecniche di machine learning e valutazione immobiliare multi-parametrica, intendendo comprendere come e quanto l'opacità nel mercato immobiliare italiano influenzi le previsioni dei valori di mercato. Per un determinato mercato (caso studio), vengono definiti i seguenti passaggi. • En esta primera fase se actualiza un software de rastreo (escaneado) automatizado en el lenguaje informático Python®. Un rastreador web y un software para abrir y analizar el contenido de la página web de forma estructurada y automatizada. Si creas un rastreador web

nell'ambito della presente ricerca al fine di análisis específico siti web di vendita immobiliare in Italia y scaricare automaticamente gli pedir precio delle proprietà attualmente en venta.

Contestualmente, el precio de venta también incluye algunas características descriptivas de los muebles que vengono raccolte tramite el medesimo web crawler. Grazie a este procedimiento è possibile ottenere agilmente migliaia di dati. Ogni dato (osservazione) rappresenta un immobile en vendita di cui si conoscono il prezzo di offerta e alcune características descriptivas.

La base de datos, asociada con el rastreador web para un mercado de datos, generalmente se denomina DBcrawler.

- En la segunda fase del proceso, viene activada una red neuronal artificial (o red neuronal artificial) a partir de DBcrawler, y en consecuencia denominada

ANNcrawler, con lo scopo di prevedere il valore di un immobile in funzione delle sue características all'interno del dato mercato.

En el campo del aprendizaje automático, las RNA son sistemas computacionales en grado de aprendizaje de patrones y procedimientos (Cetkovic et al., 2018). Le reti neurali sono costituite da neuroni artificiali e sinapsi artificiali: i neuroni sono le unità computazionali, mentre la sinapsi collegano i neuroni tra di loro. Las neuronas artificiales están organizadas en más capas separadas de modo que las neuronas de entrada se visualizan en la capa de entrada y las neuronas de salida si el contenido en la capa de salida.

Tra di essi sono presenti vari Layer di neuroni chiamati Layer nascosti. Al utilizar un procedimiento de aprendizaje automático, una red neuronal puede «impararse» como información específica de entrada fluida de las neuronas de entrada para generar una respuesta de salida.

Una red puede essere addestrata su qualsiasi base de datos de información estructurada en forma de entrada-salida, produciendo como resultado un modelo predictivo (Pitarrello et al., 2021). En este artículo, las neuronas de entrada de ANNcrawler son las características descriptivas de los inmuebles, entre las neuronas de salida representan su correspondiente valor de mercado.

En particular, le ANN qui assumono il ruolo di una tecnica di valutazione multi-parametrica del valore di mercato. Le ANN vengono infatti utilizzate para analizar diversas características del edificio y preservar el valor del mercado. Nella funzione di previsione, ogni variabile indipendente (le características dell'edificio) contribuisce in modo diverso alla stima del prezzo (la variabile dependiente) (Simonotti, 2006). De esta manera, es posible aislar el aporte que todos los demás se aplican al precio y a la captura del tipo de error en la última hora que puede producir un diverso nivel de vida de las informaciones sobre las características del móvil.

Las columnas principales de las neuronas de entrada son llamado [Xr]crawler, con  $1 \leq r \leq R$ , dove R y el número total de servidores contenidos en la base de datos. La neurona de salida es un indicador de una columna y una línea indicadora del rastreador [Y\_forecast]. [Y\_forecast] rastreador también puede estar representado como [Y\_forecast] rastreador =  $f([Xr]crawler)$ , además de una función del mismo rastreador [Xr]crawler Definida attraverso i Layer nascosti.

Cuando una red viene addestrata (entrenamiento) su base de datos que se encuentra en la relación de entrada a salida, la red «impara» como los datos de entrada son correlativos al resultado correspondiente. Di conseguenza la rete ne costruisce un modello previsionale. Durante el entrenamiento de una base de datos, de hecho, la red modifica iterativamente sus parámetros libres, vale a dire i pesi (w) y sesgos (b) de la red, hasta identificar el mejor modelo previo.

Alla scala del singolo neurone, durante el entrenamiento, ogni zesimo neurone ottiene uno o más input numerici, qui chiamati  $xz, u$ , con  $1 \leq u \leq U$ , dove U è il numero di sinapsi artificiali che entrano nel neurone. Quindi le in-



formazioni ricevute vengono tutte combinate all'interno del zesimo neurone attraverso una funzione definita da un peso ( $wz,u$ ) e da un sesgo ( $bz$ ), di modo che venga un prodotto di salda numerico ( $Yz$ ). En este punto, una función de activación ( $z$ ) transforma el valor de la neurona en un valor de respuesta en forma de somma ponderata, así que definitivamente se cruza con la ecuación 1:

$$\text{neurona zesimo, } Yz = z\left(\sum u^{\text{ud}} = 1[wz, u * xz, u] + bz\right) \quad (1)$$

Durante el agregado de la red de pesos y sesgos vengono quindi modificati iterativamente fino al trapo del mejor modelo de previsión o, en otras palabras, fino a quando  $[Y\_forecast]_{crawler}$  diventa il più similar possibile a  $[Y\_expected]_{crawler}$ . En este caso, el rastreador  $[Y\_expected]$  es el valor objetivo, además del valor del contenido del mercado en la base de datos de entrenamiento ( $DBcrawler$ ), y el rastreador  $[Y\_forecast]$ , definido antes de la fecha, es el producto valorado de ANNcrawler. La diferencia entre el valor objetivo y la previsión produce una señal de error denominado  $[err]_{crawler}$  en la ecuación 2:

$$[err]_{\text{rastreador}} = \text{Rastreador } [Y\_expected] - \text{Rastreador } [Y\_forecast]. \quad (2)$$

Durante la adición de ANNcrawler a través de DBcrawler, el objetivo es eliminar al mínimo algunas señales de error de producto.

- En la tercera fase, los autores intentan verificar el nivel de transmisión de información recopilada automáticamente mediante el rastreador web. De hecho, oltre all'incertezza propria dell'uso degli pedir precio nelle valutazioni immobiliari, è possibile verificare che gli annunci di vendita contengono anche false dichiarazioni, informazioni errate o dati incompleti.

Poiché il presente articolo mira a valutare come e quanto l'opacità nelle informazioni influenzi la corretta stima del valore di mercato dell'immobile, nella terza fase di sviluppo della ricerca, un set di controllo dei dati viene nuovamente raccolto, ma esta volta a mano.

Per quanto sia estremamente dispendiosa in termini di tempo, este procedimiento de raccolta manuale rappresenta l'unico modo per verificare la correttezza di tutte le informazioni contenute negli annunci di vendita, in modo tale da escludere gli annunci incompleti, i cui dati non possono essere verificati, completee le informazioni mancanti ma verificabili e correggere i dati errati. Este conjunto de control raccolto manualmente viene denominado DBhand.

- En cuarta fase, el ANNcrawler precedentemente sviluppato viene probado en DBhand. Esta volta, i neuroni di input sono le caratteristiche degli immobili raccolte e corrette a mano, quindi denomina  $[Xr]_{hand}$ , entre il neurone di output è la conseguente previ-

sione di un valore di mercato  $[Y\_forecast]_{mano}$ . Il valore di mercato atteso, o valore target, è il prezzo indicato negli annunci di vendita, denominado  $[Y\_expected]_{hand}$ . Claro,  $[Y\_expected]_{mano} = [Y\_expected]_{rastreador}$ . La diferencia entre el valor del mercado atteso y su previsión representa una estimación del error del producto a causa de la opacidad informativa de contenidos negli annunci online (opacidad del mercado), como se ilustra en la ecuación 3:

$$[err]_{\text{mano}} = [Y\_expected]_{\text{mano}} / \text{rastreador} - [Y\_forecast]_{\text{mano}} \quad (3)$$

## 4. TRE CASI STUDIO RAPPRESENTATIVI NEL NORD ITALIA

### 4.1 Descargar la base de datos: el rastreador web

Per questa ricerca è stato scelto di analizzare l'opacità informativa in tre diversi mercati italiani, ovvero i mercati immobiliari di:

- Bologna;
- Padua;
- Treviso.

Secondo l'opinione degli Autori, queste tre città potrebbero essere casi di studio rappresentativi perché incarnano tre differenti dimensioni di mercato, dove per dimensione di mercato qui ci si riferisce sia alle dimensioni stesse della città in analisi, sia al numero di elementi (osservazioni) che costituiscono la domanda e l'offerta, sia al volume totale delle transazioni. Además, sebbene sia chiaro che ogni mercato immobiliare rappresenta un caso specifico y único, nessuna di queste tre città descriva un caso fortemente a sé stante come, ad esempio, potrebbero essere i mercati di Venezia, Roma o Milano. I risultati possono quindi essere significativi anche per altri mercati immobiliari simili in Italia.

El primer problema es captar de manera realista los datos y la información necesaria para detectar las ANN. Come già introdotto nel Capitolo 3 è stato sviluppato un software di rastreo automatizzato

Utilice el lenguaje de programación Python para analizar con precisión los propietarios de inmuebles en venta, Bologna, Padua y Treviso en la web de venta de modo que se cargue automáticamente el precio de venta y las características.

La arroz en línea del rastreador web debe ser objetivo a través de la definición de tres diversos dominios de cerca, paloma ogni dominio è específico para una ciudad. I tre domini comprendono tutti gli immobili residenziali in vendita a Bologna, Padua y Treviso. Si elige salir de este estudio como propietario no residencial, como propietario comercial y en dirección, así como también sobre los puestos locales en cuestión. Per quanto riguarda la ti-

## Identificar la transferencia informativa en el mercado inmobiliario italiano: una aplicación de aprendizaje automático

pologia edilizia, sono incluse nel dominio di ricerca sia le nuove costruzioni che gli edifici esistenti, quali appartamenti, soffitte, villette a schiera e ville pluri/bi/uni-familiari.

Para el primer dominio de Ricerca, todas las 10 áreas de Bologna se incluyen en el procedimiento de descarga, mientras que el segundo dominio contiene todas las 14 áreas en cui è subdiviso il Comune di Padova. Infine, le 7 aree di Trevi sono altresì tutte comprese nel terzo dominio di ricerca.

Para permitir que el rastreador web extraiga la información de cada anuncio de venta, es necesario insertar todo el código Python en el Localizador uniforme de recursos (URL) de cada anuncio bajo forma de dirección web, como «https://...».

En este momento, claro, abre todos los anuncios de venta online para copiar e incluir la URL en el interior del código Python es una operación manual impracticable, el objetivo y la captura como resultado de la URL en modo automático. È stato notato che ogni annuncio di

vendita mostra un URL que es el formato de la combinación de la URL de la página de inicio más un número de serie, la página de inicio puede ser definitiva como el elemento de todos los anuncios de venta en el sitio disponible en línea en el interior de los datos del dominio.

Por lo tanto, el rastreador web está programado para rastrear primero la URL de la página de inicio, puede identificar los números de la serie de anuncios en ese contenido y crear automáticamente las URL correspondientes.

Sucesivamente, está implementada la librería Py-thon «Beautiful Soup» para analizar y analizar todos los archivos.

página HTML de los anuncios en venta. «Beautiful Soup» es un paquete diseñado por Leonard Richardson a principios de análisis fino de documentos en formato HTML. Con el contenido de esta librería, es posible extraer datos e información de las pruebas HTML, creando así un albero de análisis para cada página web identificada. Ahora podemos construir en Python una clase de objetos y funciones para definir la información de acceso a cada anuncio. La clase utilizada en ciascun dominio di ricerca è illustrata in Tabella 1.

La clase de objetos y funciones sopra definiti è stata determinata sia in base a un análisis específico de la información disponible contenute nei siti di vendita immobiliare, sia in base a todas las características de los edificios más comunmente utilizados en el procedimiento de estimación multiparamétrica del valor . di mercato (Feng y Zhu, 2017; Wang y Xu, 2018). Tra gli attributi di stima si considerano caratteristiche strutturali/fisiche dell'edificio, qualità del quartiere e localizzazione.

A continuación, todos los análisis de las páginas HTML de los anuncios de venta, ahora podemos implementar en Python la librería de análisis de los datos «Pandas». Desarrollado por Wes McKinney, la librería «Pandas» se utiliza para guardar un archivo en formato .xls a continuación de un procedimiento de escaneo de sitio web con el alcance de visualización de datos y la información estratificada en forma de tabla. Ogni riga della tabella rappresenta una diversa osservazione, mentre le colonne indicano gli elementi della classe scaricati per ogni ogni osservazione. La base de datos de Bologna, chiamata DB(B)crawler, presenta 2.455 reservas, en la base de datos descargada en Padua, denominada DB(P) rastreador, tenemos 2.884 reservas, entre la base de datos de Treviso, definitivamente como DB(T)crawler, muestra 1.473 observaciones.

Tabla 1 - Clases de objetos y funciones

Elemento o función de la clase	Udm	Elemento o función de la clase	Udm	Elemento o función de la clase	Udm
IDENTIFICACIÓN		PRECIO		ascender	binario
URL web	texto	precio	€	Giardino privado	binario
Numero anuncio	número	Sorber. comercial	m <sup>2</sup>	Garaje privado	binario
código identificativo	número	Precio unitario	€/m <sup>2</sup>	comuna de Giardino	binario
LOCALIZACIÓN		CARACTERÍSTICAS		parcheio	binario
zona	texto	norte. bolso	número	Cantina	binario
latitudina	coordenada	norte. cámara	número	Terraza	binario
longitudina	coordenada	Piano mín.	número	Domótica	binario
TIPOLOGIA EDILIZIA		Piano máx.	número	Riscaldamento centralizado binario	
Apartamento	binario	último piano	binario	Impianto fotovoltaico	binario
soffita	binario	CONDICIONES		Ventilación mecánica	binario
Casa a schiera	binario	Clase energética	número de(1 a 10)	Aria condizionata	binario
Villa plurifamiliar	binario	manutenición	número (del 1 al 4)	fibra otica	binario
Villa bifamiliar	binario	Año de construcción	numero (año)	Camino	binario
villa individual	binario	ATRIBUTOS		alarmarme	binario

[

4.2 Base de datos Pulizia dei: rimozione delle  
observaciones incompletas y degli valores atípicos

Después del procedimiento de descarga, es necesario limpiar la base de datos de capacitación de las ANN con observaciones incompletas o evidentemente erróneas. Come prima cosa sono stati esclusi gli annunci incompleti. In un secondo momento sono poi state escluse le osservazioni contenenti errori evidenti o valori anomali (outlier), ven, ad esempio, quando contenenti un prezzo di vendita nullo o una superficie commerciale nulla. En particular, las observaciones incompletas/ erradas se ilustran en la Tabla 2, definidas para cada ciudad y elemento de la clase.

Pertanto, le percentuali espresse nella Tabella 2 rappresentano le osservazioni che sono state eliminate dal campionamento perché i dati non erano completi (o corretti) in tutte le classi di oggetti. Por ejemplo, la «classe energética» y el «año de construcción» representan el más alto percentual de observaciones perse a causa di informazioni mancanti.

Pertanto, la base de datos de formación descargada a través del rastreador web hanno dovuto essere ridotti, rispettivamente, a 1.665, 2.122 y 867 elementos.

5. LE TRE RETI NEURALI

5.1 El entrenamiento transmite el  
algoritmo de optimización “Cuco”

La extracción de tres bases de datos (DBscrawler) desde las observaciones estadísticamente no significativas es posible ejecutar la formación de los tres corresponsales de ANNscrawler. Configuré las mejoras de entrenamiento para agregar los nervios de los nervios a una etapa determinada, seleccionando casualmente el 60 % de las operaciones de almacenamiento de DBcrawler para cada ciudad. Configuré la selección. Sono definiti prendendo in modo casuale un altro 20% delle istanze rimanenti, mentre i dati residui creano i rispettivi set di test. Sono state valutate tre diverse proporzioni di suddivisione dei base de datos: 80%-10%-10%, 70%-15%-15% o 60%20%20%. La proporción de división definitiva es estadísticamente selecta al número de observaciones y entradas presentes en el conjunto de datos y en los modelos de red neuronal. En este caso, la base de datos sono abbastanza ridotti, todas las neuronas de entrada rendono il modello molto complesso. Esto ha reso fondamentale mantenere un número adecuadamente suficiente de osservazioni nei set di selezione e test.

El procedimiento de entrenamiento está desarrollado en el código Python y se implementa por separado para cada ciudad. Establecí di addestramento vengono inizialmente utilizzati per generare diversi modelli di ANN. A continuación, estos modelos son implementados en su conjunto de selección al fin de identificación cual ANN realiza mejor desempeño que su último conjunto. El conjunto de pruebas se utiliza infinitamente para calcular el error en las predicciones.

El proceso de formación viene eseguito all'interno di un procedura di ottimizzazione che consente di testare di-

Tabla 2 - Porcentaje de anuncios persi por clase

clase	Porcentual de datos eliminados		
	Bolonia	Padua	Treviso
zona	1,37%	0,43%	5,06%
latitudina	1,34%	2,77%	5,50%
longitudina	1,55%	0,73%	5,57%
Apartamento	0,00%	0,00%	0,00%
soffita	0,00%	0,00%	0,00%
Villa plurifamiliar	0,00%	0,00%	0,00%
villa individual	0,00%	0,00%	0,00%
Casa a schiera	0,00%	0,00%	0,00%
Villa bifamiliar	0,00%	0,00%	0,00%
precio	2,44%	1,80%	2,65%
Sorber. comercial	1,96%	1,32%	0,41%
norte. bolso	2,61%	3,16%	3,73%
norte. cámara	4,56%	3,47%	3,39%
Piano mín.	2,32%	2,70%	5,84%
Piano máx.	2,32%	2,70%	5,84%
último piano	1,18%	0,45%	0,20%
Clase energética	24,52%	22,33%	28,45%
manutención	4,56%	6,14%	7,06%
Año de construcción	19,85%	28,54%	21,90%
ascender	1,18%	0,45%	0,20%
Giardino privado	1,18%	0,45%	0,20%
Garaje privado	1,18%	0,45%	0,20%
comuna de Giardino	1,18%	0,45%	0,20%
parcheio	1,18%	0,45%	0,20%
Cantina	1,18%	0,45%	0,20%
Terraza	1,18%	0,45%	0,20%
Domótica	1,18%	0,45%	0,20%
Riscaldamento centralizado	1,18%	0,45%	0,20%
Impianto fotovoltaico	1,18%	0,45%	0,20%
Ventilación mecánica	1,18%	0,45%	0,20%
Aria condicionada	1,18%	0,45%	0,20%
fibra optica	1,18%	0,45%	0,20%
Camino	1,18%	0,45%	0,20%
alarmarme	1,18%	0,45%	0,20%

Versi modelli di rete neurale per ridurre al minimo l'errore sulle previsioni. El algoritmo de optimización utilizado durante el entrenamiento delle reti es el algoritmo de optimización “Cuco” (Chiroma et al., 2017; Mareli e Twala, 2018). Esto es un algoritmo de optimización inspirado en la naturaleza.

[

## Identificar la transferencia informativa en el mercado inmobiliario italiano: una aplicación de aprendizaje automático

che identifica e confronta tutte le possibili architetture delle ANNs che hanno un punto di ottimo relativo (cioè l'errore minimo) fino ad identificare quella che presenta l'ottimo globale. La ottimización del algoritmo "Cuco" es infatti adatta a risolvere i problemi di ottimiz-zazione globale poichè impiega il processo casuale dei "random walks".

Por lo general, la red neural sviluppata de Bologna se indica como ANN(B)crawler, como Padua.

ANN(P)crawler, y por Treviso se llama ANN(T)crawler. I risultati delle ANNs sviluppate sono presentati in Tabella 3. Los errores en los productos en el conjunto de pruebas son del 9,50% para Bologna, del 11,75% para Padua y del 13,55% para Treviso.

Gli errori sono circa dello stesso ordine di grandezza. Sin embargo, podemos notar una correlación inversa de riesgos en todas las dimensiones del mercado: más grande/activo y el mercado menor y el error, y viceversa.

Tabla 3 - Caratteristiche delle ANN

NN	N. di entradas	Producción	Numeros de capas nascosti	Numero de neuronas por capa nascosto	Funciones de activación	Escalarización	Capacitación
rastreador ANN(B)	32	Valor de mercado	7	184	ReLU	mínimo máximo	error medio cuadrado
rastreador ANN(P)	32	Valor de mercado	7	144	ReLU	mínimo máximo	error medio cuadrado
rastreador ANN(T)	32	Valor de mercado	7	144	ReLU	mínimo máximo	error medio cuadrado

### 5.2 Verificación del transparencia informativa

Esta sección está dedicada a todos los usos del ANNscrawler sviluppate en precedenza para discutir como el traspa-renza del mercado (o mejor su opacidad) influye en la afidabilidad del tiempo en el mercado inmobiliario italiano.

A tale scopo, i tre Database sono stati raccolti nuova-mente, ma esta seconda volta a mano, per verificare la correttezza delle informazioni che in precedenza erano state scaricate automaticamente dal web crawler e utiliz-zate per costruire le reti. I base de datos raccolti a mano sono chiamati DB(B)hand por Bologna, DB(P)hand por Padua, y DB(T)hand por Treviso.

Obviamente, para definir los tres dominios de ricerca web sono stati utilizzati gli stessi criteri di selezione già utilizzati según la precedente ricerca online. I domini si limitano così a cercare gli immobili residenziali in vendita (esclusi gli af-fitti e altri tipi di contratto), comprendiendo sia edifici esi-stenti che nuove costruzioni.

Durante esta segunda recopilación de datos viene verificada la corrección de toda la información, y todos los inmuebles de cui no es posible verificar las características de vengono directamente escluidas de DBshand.

Certamente este método de recopilación de datos es un proceso muy completo y disponible en términos de tiempo, ma, allo stesso tempo, el único posible para probar y verificar la transferencia del mercado, la veridicidad de los datos y la disponibilidad de información.

La base de datos DBshand sono stati implementati atraverse le ANNscrawler in modo che le caratteristiche corrette a mano degli immobili ne costituiscano i neuroni di input per poi produrre una nuova stima del loro valore di mer-cato. Come previsto dagli Autori, l'errore prodotto su que-ste ultime previsioni è superiore rispetto a prima. El error medio por Bologna es del 19,45%, por Padua es del 25,57%

e por Treviso è del 26,97%. Este aumento de los errores en las previsiones y la opacidad del mercado o, en otras palabras, todas las informaciones erradas informan negli annunci. Más grande es el mercado inmobiliario, menor y el porcentaje de error medio: los mercados más pequeños son más pequeños trasparenza nei dati descritti dagli annunci immobiliari.

La localización si revela essere la fuente principal de opa-cità. Solo alcuni tra gli annunci controllati mostrano la posizione esatta dell'edificio, entre la mayor parte de essi sono posizionati in una strada o quartiere non corretti.

En general este problema si è rivelato più marcato a Padova e Treviso rispetto Bologna. Otros anuncios, además, son cosas propias que excluyen estos análisis porque no cuentan con información suficiente (imágenes o descripciones) para permitir esa localización de la posición real del inmóvil. Certamente, le ragioni alla base di esta mancanza di trasparenza nella localizzazione dell'immo-bile possono risiedere sia nella Privacy richiesta dai pro-prietari, sia nella strategia Commerciale adottata dall'agen-zia Immobiliare. Infatti, diviene necessaria una interme-diazione qualora un potenziale acquirente non sia in grado di identificare la posizione dell'immobile di interesse. Ciò nonostante, la mancanza di chiarezza riguardo la localizzazione rimane eccessiva e fuorviante, e diversi annunci especificano una posizione degli immobili che si è rivelata completamente sbagliata, confondendo persino aree centrali, semi-centrali e periferiche.

Una fuente de opacidad muy importante es si revela essere il li-vello di manutenzione poichè, spesso, non corresponde alle altre informazioni reportate negli annunci stessi come, ad esempio, la classe energetica, el año de construcción o los imperativos disponibles. En general, las condiciones de ma-nutención non sono coerenti con il vero stato in cui versa l'immobile. In alcuni casi le immagini mostrano li-velli di manutenzione chiaramente molto peggiori di

[ quanto dichiarato negli annunci. Questa valutazione rimane comunque qualitativa e soggettiva, e gli Autori hanno ritenuto di modificare le condizioni di manutenzione dichiarate solo quando indubbiamente diverse dalle immagini fornite.

Otros errores específicos de contenido negli annunci riguardano la identificación y la definición de la tipología edilizia, la disponibilidad de una cantina, el livello di piano, la presenza de un garaje y de un jardín privado/comune. En particular, gli annunci o omettono queste informazioni oppure non forniscono dati coerenti. Gli annunci, infatti, che sono costituiti sia una descrizione verbale che da una sintesi in tabella, presentano alcune informazioni nella descrizione nettamente in contrasto rispetto quelle inserite in tabella.

Ancora, la definizione di attico rimane vaga e manca di trasparenza. La parola italiana utilizzata negli annunci per indicare un attico dovrebbe essere utilizzata per gli appartamenti di lusso e posti all'ultimo piano di un edificio. Sin embargo, en algunos anuncios, el término «ático» è anche utilizzato per appartamenti e soffitte di categoria normale o economica. A causa dell'incertezza data da este para-metro, gli Autori hanno deciso di segnalare solo lo stato di «ultimo piano», senza alcun riferimento al livello qualitativo dell'immobile.

Otros errores riguardano gli impianti: l'aria condizionata o gli impianti di ventilazione meccanica sono spesso dichiarati presenti nella casa anche quando ne esiste only una predisposizione di tubi e condutture per una futura installazione ipotetica.

También hay otra zona de opacidad de la información que no se correlaciona directamente con las características descriptivas de los edificios. Gli annunci online subiscono un alto tasso di modifica/decorrenza. Di conseguenza, molti annunci sono stati rimossi, reintegrati o modificati in un arco temporale di poche settimane. Ciò ha reso necessario aggiornamenti frequenti e nuove download della base di dati, e per questo motivo il rastreador web automatizzato e rastreado in el ámbito de esta arrozca si è rivelato estremamente utile. Sin embargo, también los controles manuales devono essere eseguiti più volte, aumentando significativamente el cargo de trabajo total.

Otra cuestión es proteger la distribución asimétrica de las informaciones. Por ejemplo, los anuncios para los edificios de nueva construcción son generalmente más ricos en datos de edificios antiguos. Como consecuencia, es necesario descartar durante el control manual un número mayor de anuncios para el edificio vetusti a causa del mantenimiento de la información crucial. Por este motivo, la base de datos presenta una frecuencia mayor de anuncios de ofertas de edificios nuevos.

## 6. DEFINIR LA IMPORTANCIA DELLE VARIABILI

En el párrafo anterior, sono state discusse le fonti primarie di opacità. Sin embargo, è da considerare che non tutti gli errori contenuti negli annunci producono lo stesso im-

patto sulla stima del valore di mercato. La opacità in alcune informazioni es mucho más significativa de otras. Por este motivo, se puede utilizar para determinar la calidad de los parámetros de entrada y mostrar el mayor impacto en el valor del mercado a través de un análisis de importancia de características (o análisis de la influencia de las variables).

Tra gli approcci che aiutano a calcolare l'impatto delle variabili sull'output ci sono i metodi basados en filtros, i métodos Wrapper y gli métodos integrados (Tatwani e Kumar, 2019).

Los métodos basados en filtros se basan en estadísticas univariadas, como el coeficiente de correlación de Pearson, la prueba de cuadro, la puntuación de Fisher, el umbral de varianza, el índice de dispersión o la diferencia absoluta de medias. Los métodos Wrapper imponen la selección de un conjunto de características, che a guisa di un problema di ricerca (Ghosh et al., 2020; Suresh e Narayanan, 2019; Yassi e Moattar, 2014), quali la Forward Feature Selection, la Backward Feature Elimination, l'Selección exhaustiva de funciones o la Eliminación recursiva de funciones. Además, los métodos integrados combinan las características de los métodos Filter y Wrapper, como la regularización LASSO y el bosque aleatorio.

En este artículo se evalúa la influencia de las características a través de la metodología del Random Forest (RF) o foresta casual. È stato scelto este approccio perché gli Los métodos integrados son altamente precisos y son excelentes propiedades de generalización (Siham et al., 2021).

Una selva casual y un formato de clasificación particular de un conjunto de decisiones (clasificación simple) (Ugolini, 2014), dove un albero decisionale, nel campo dell'informatica, è una struttura dati costituita da nodi e archi. Un albero decisionale viene letto dall'alto verso il basso. I nodi dell'albero sono gli elementi che contengono le informazioni, mentre gli archi sono le connessioni tra i nodi. Il nodo iniziale è la radice e non ha archi di input, mentre i nodi terminali sono chiamati foglie e non mostrano archi in uscita.

Ogni albero decisionale, durante un procedimento di foresta casual, viene costruito (cioè addestrato) a partire da un sottoinsieme definito in maniera casuale del set di training. En

este caso, se establecen de formación sono le banche dati di informazioni descrittive degli edifici raccolte rispettivamente per Bologna, Padova e Treviso. Ogni albero decisionale è costruito su un'estrazione casuale delle caratteristiche analizzate (cioè le caratteristiche degli immobili). Esta

casualidad en la selección de las características y las observaciones es una parte fundamental durante la construcción de los clasificadores y tiene el alcance de aumentar la heterogeneidad y disminuir la correlación entre las variables.

Para definir la importancia de todas las características, es necesario medir cuanto disminuya la impureza de todas las características durante el entrenamiento. De hecho, más una variable disminuye la propia impureza, más significativamente resulta essere quella variable. En la clasificación (variables discretas), la impureza y los datos de la impureza de Gini o del aumento/riducción de entropía. Nelle regressioni (las variables continúan),



## Identificar la transferencia informativa en el mercado inmobiliario italiano: una aplicación de aprendizaje automático

invece, l'impurità è quantificata attraverso dalla varianza. È stata definita una matrice colonna in cui le assi x rappresentano le caratteristiche degli immobili (le variabili) e l'asse y mostra l'obiettivo (il valore di mercato). Esta vez utilizé en Python la librería "Numpy" para seguir el regreso de RF. Il regressore-RF è in grado di calcolare i coefficienti di importancianza per ogni caratteristica. Este estado se utiliza como conjunto de adición del 70 % de las reservas, mientras que el restante 30 % se aplica como conjunto de prueba. Durante el procedimiento RF sono stati costruiti 2.000 alberi e il umbral impostato è pari allo 0,75 del valor medio de los coeficientes de importancia calculati. En este caso, la diminuzione dell'impurità di ogni caratteristica viene valutata come la media delle diminuzioni date da ciascun albero costituente la foresta. De este modo viene estimada la importancia final de ciascuna variable. Los coeficientes de importancia calculados del regresor-RF sono ilustrados en la Tabla 4.

Supponiamo che un'informazione errata negli annunci riguardi i dati di maggiore impatto, come la latitudine e la longitudine (quindi la localizzazione), il livello di manutenzione o la superficie commerciale. En tal caso, se producirá un error en la previsión del valor del mercado. Al contrario, las variaciones menores son los impedimentos y la tecnología como la climatización, la ventilación mecánica, el encendido, el ascensor o la domótica. Tra le variabili meno impattanti ci sono anche il piano interrato, il giardino condominiale e il livello di piano.

### 7. DISCUSIÓN Y CONCLUSIÓN

Este trabajo ha integrado el análisis de mercado con técnicas multiparamétricas de estimación del valor de mercado, programación informática y procedimiento de aprendizaje automático para comprender la opacidad del mercado en Italia.

En primer lugar, un rastreador web automatizado, desarrollado en el lenguaje Python, ha ayudado a rastrear rápidamente una nota de cantidad de datos descriptivos riguardo gli immobili en venta en Bologna, Padua y Treviso. La base de estas tres bases de datos son administración de estado en Python tres corresponsales neuronales artificiales al fin de preservar el valor del mercado de una propiedad en la función de descripción descriptiva de 32 quilates de entrada, tra cui, la localización, el vivero de manutención, gli impianti, la tipologia edilizia, la presenza di terrazza, garage o giardino. A este punto i tre base de datos sono stati raccolti una segunda volta ma-nualmente. Este procedimiento ha permitido controlar la veracidad de alguna información indicada negli annunci. Le informazioni errate sono state corrette, mentre le osservazioni non verificabili sono state escluse. Questi tre base de datos «corretti» sono stati implementati nelle reti neurali artificiali sviluppate in precedenza: l'errore prodotto sulle previsioni rappresenta l'errore nella stima del valore di mercato dovuto all'opacità del mercato stesso (o, in altre parole, l'errore dovuto alle informazioni non corrette contenute negli annunci). Infine, è stata eseguita un'ana-

Tabla 4 - Coeficientes de importancia RF

Elemento o función de la clase	coeficientes RF		
	Bolonia	Padua	Treviso
variable	%	%	%
latitudina	22,43%	19,69%	15,38%
longitudina	29,30%	19,52%	13,38%
Tipología	0,75%	0,72%	1,05%
Sorber. comercial	8,17%	7,35%	7,71%
norte. bolso	0,87%	0,93%	1,20%
norte. cámara	1,50%	1,12%	0,97%
Piano mín.	2,03%	2,07%	4,02%
Piano máx.	1,53%	1,63%	1,43%
último piano	0,61%	0,40%	0,37%
Clase energética	3,19%	3,31%	6,99%
manutención	14,76%	20,19%	25,80%
Año de construcción	7,55%	15,46%	10,03%
ascender	0,62%	1,00%	0,49%
Giardino privado	0,65%	0,35%	0,34%
Garaje privado	0,50%	0,53%	0,99%
comuna de Giardino	0,53%	1,01%	4,11%
parcheio	0,61%	0,59%	0,51%
Cantina	0,81%	0,78%	0,52%
Terraza	0,60%	0,53%	0,78%
Domótica	0,15%	0,09%	0,12%
Riscaldamento centralizado	0,62%	0,40%	1,34%
Impianto fotovoltaico	0,08%	0,22%	0,29%
Ventilación mecánica	0,00%	0,16%	0,23%
Aria condicionada	0,46%	0,29%	0,23%
fibra optica	0,59%	0,54%	0,50%
Camino	0,50%	0,45%	0,59%
alarmarme	0,58%	0,68%	0,64%

lisi dell'influenza delle caratteristiche descriptive basata sulla metodologia Random Forest.

En conclusión, esta arroz puede ayudar a capire como y cuanto la opacidad del mercado en Italia influye en su affidabilità delle stime immobiliari. Las redes neuronales artificiales son procedimientos estadísticos eficaces de previsión y modelos de redes neuronales en el grado de descripción precisa de la relación entrada-salida. El uso de un modelo de red neuronal para confrontar los resultados de una base de datos opaco rispetto a una base de datos «corretto» ha por-tato a determinante quanto un valutatore possa essere in-dotto a sbagliare la stima del più probabile valore di mer-cato di un immobile solo a causa della mancanza di tra-sparenza informativa.

[

Además, este análisis multiparamétrico también permite identificar las variables que más influyen en la estimación del valor del mercado. Esto es un aspecto crucial en cuanto a los datos informativos del edificio, es necesario controlarlos atentamente para valorar de modo corregido el valor del mercado inmobiliario. Claro, una mayor opacidad en las variables más impacientes portará a un error más elevado en la previsión, mientras una menor reserva en las variables menos influyentes produce un error menor en la estimación.

Al terminar esta arzo, los autores indican la necesidad de mejorar la división de datos y la información de sus propietarios inmobiliarios en Italia. No solo i prezzi di compravendita dovrebbero essere resi disponibili, ma anche le descrizioni degli immobili dovrebbero essere muy precisa y completa. Además, los anuncios de ven-dita potrebbero richiedere un livello minimo di dati da fornire prima di essere considerati pronti ad essere messi online. Ancora, le informazioni fornite negli annunci dovrebbero essere precision e corrette, soprattutto per quanto riguarda la localizzazione della proprietà. Infine, gli annunci dovrebbero essere più trasparenti: anche siti web diversi tra loro potrebbero comunque rispettare un diseño condiviso per garantire completezza e chiarezza dell'informazione.

Lo scopo è duplice: in primo luogo, ridurre l'asimmetria informativa tra venditore e acquirente per far sì che la domanda possa essere in grado di muoversi in maniera più consapevole all'interno del mercato immobiliare. En se-

condo luogo, lo scopo è aumentare la trasparenza del mercato, poiché tali dati sono utilizzati sia da società che analizzano i mercati, producendo relazioni e pubblicando prezzi, sia dai valutatori che sono talvolta costretti a basarsi sugli prezzi in quanto i prezzi di compravendita non sono disponibili. Además, el mercado inmobiliario es completo y tiene implicaciones sustanciales para el resto de la economía, todos los operadores deben asegurarse de la calidad de la información que utilizan.

La problematica affrontata in questa ricerca ruota attorno al concetto di «qualità dell'informazione». Naturalmente, una firma de mercado no puede prescindere dall'analisi approfondita di ogni dato di riferimento, che si tratti di un acquisto, una vendita o un'offerta. Además de esto, sembra poi utile concentrarsi anche sulle procedimientos di stima oltre che sul tipo di dato: una metodologia appropriata consente di approcciare in modo professionale e costruttivo anche informazioni spurie, traendone intuizioni significativa.

Per gli ulteriori sviluppi di esta ricerca, gli Autori intendono applicare periodicamente la metodologia adottata in altri mercati immobiliari italiani al fine di mappare i diversi livelli di opacità nei mercati e per capire se vi sia un'evoluzione nel tempo per quanto riguarda l'accesso a toda la información. En particular, potrebbe essere interessante comprendere se le dinamiche che si sono verificate negli ultimi anni (pandemia di Covid-19, guerra in Ucraina, crisi energetica) possono, in qualche modo, impattare non solo sulle dinamiche e sui prezzi degli immobili ma anche sul livello di trasparenza o opacità.

\* Laura Gabrielli, Dipartimento di Culture del Progetto, IUAV Università di Venezia correo electrónico: laura.gabrielli@iuav.it  
\*\* Aurora Greta Ruggeri, Dipartimento di Culture del Progetto, IUAV Università di Venezia correo electrónico: aurora.ruggeri@iuav.it  
\*\*\* Massimiliano Scarpa, Dipartimento di Culture del Progetto, IUAV Università di Venezia correo electrónico: massimiliano.scarpa@iuav.it

## nomenclatura

[err]crawler - error en ANNcrawler

[err]mano - error de DBhand

[Xr]crawler - vettore colonna dei neuroni di input di ANNcrawler

[Xr]hand - vettore colonna dei neuroni di input da DBhand

[Y\_expected]crawler - valor objetivo en ANNcrawler

[Y\_expected]hand - valor objetivo de DBhand

[Y\_forecast]crawler vettore una a riga e una colonna del neurone de salida ANNcrawler

[Y\_forecast]mano - neurone de salida de DBhand

ANN(B)crawler, - ANN desarrollado por DBcrawler para Bologna

ANN(P)crawler, - ANN desarrollado por DBcrawler para Padua

ANN(T)crawler - ANN desarrollado por DBcrawler para Treviso

ANN/ANN - Red neuronal artificial/Neural artificial

Redes

ANNcrawler - Red neuronal artificial desarrollada

DBcrawler

bz - sesgo

DB(B)crawler - rastreador web raccolto tramite de bases de datos por Bologna

DB(B)hand - base de datos raccolto a mano per Bologna

DB(P)crawler - rastreador web raccolto tramite de bases de datos por Padua

DB(P)hand - base de datos raccolto a mano per Padova

DB(T)crawler - rastreador web raccolto tramite de bases de datos por Treviso

## Identificar la transferencia informativa en el mercado inmobiliario italiano: una aplicación de aprendizaje automático

DB(T)hand - base de datos raccolto a mano per Treviso

DBCrawler - rastreador web de base de datos raccolto tramite

DBhand - base de datos raccolto a mano

RF - Bosque aleatorio

U - numero di sinapsi artificiali entranteti nel neurone wz,u - funzione dei pesi nel neurone

xz,u - entrada numérica en la neurona

Yz - salida numérica en la neurona

z - Función de activación de la neurona.

### Bibliografía

AKERLOFF GA, El mercado de los "limones": la incertidumbre de la calidad y el mecanismo del mercado, *The Quarterly Journal of Economics*, 84(3), 1970, págs. 488–500.

ANGLIN PM, RUTHERFORD R. AND SPRINGER TM, La compensación entre el precio de venta de propiedades residenciales y el tiempo en el mercado: el impacto de la fijación de precios, *The Journal of Real Estate Finance and Economics*, vol. 26, núm. 1, 2003, págs. 95–111.

ARNOTT R., Teoría Económica y Vivienda. En *Handbook of Regional and Urban Economics*, editado por E. Mills, Londres: Elsevier, 1987, págs. 959–988.

BERACHA E. y SEILER MJ, El efecto de la estrategia de precios de cotización en los precios de venta de transacciones, *The Journal of Real Estate Finance and Economics*, vol. 49, núm. 2, 2014, págs. 237–255.

C' ETKOVIC' J., LAKIC' S., LAZAREVSKA M., ŽARKOVIC' M., VUJOEVIC' S., CVIJOVIC' J. AND GOGIC' M., Evaluación del valor de mercado inmobiliario en el mercado europeo mediante la aplicación de redes neuronales artificiales, *Complejidad*, vol. 2018, artículo ID 1472957, 2018, p. 10.

CHIROMA H., HERAWAN T., FISTER I., FISTER I., ABDULKAREEM S., SHUIB L., HAMZA MF et al., Computación bioinspirada: desarrollo reciente sobre las modificaciones del algoritmo de búsqueda del cuco, *Computación suave aplicada*, vol. 61, 2017, págs. 149–173.

CURTO R., FREGONARA E. Y SEMERARO P., Comportamiento de cotización en el mercado inmobiliario italiano, *Revista Internacional de Análisis y Mercados de la Vivienda*, vol. 8, núm. 1, 2015, págs. 97–117.

FARZANEGAN MR Y FEREIDOUNI HG, ¿Importa la transparencia inmobiliaria para las inversiones inmobiliarias extranjeras?, *Revista Internacional de Gestión Estratégica de Propiedades*, vol. 18, núm. 4, 2014, págs. 317–331.

FARZANEGAN MR, GHOLIPOUR HF, ¿Importa la transparencia inmobiliaria para las inversiones inmobiliarias extranjeras? En t. J. Estrategia. Gestión de prop. 18 (4), 2014, págs. 317–331.

FENG J. Y ZHU J., Modelo de regresión no lineal y análisis de opciones del precio de bienes raíces, *Dalian Ligong Daxue Xuebao/Revista de la Universidad Tecnológica de Dalian*, vol. 57, núm. 5, 2017, págs. 545–550.

FORTE C. Y DE ROSSI B., *Principi Di Economia Ed Estimo*, Etas., Milán, 1974.

FRANCÉS N., Precios de mercado. Una investigación sobre el uso de pruebas comparables en la valoración de propiedades, *TEGoVA*

Asociación del Grupo Europeo de Tasadores, junio de 2020.

FRENCH N., CROSBY N. Y THORNE C., Fijación de precios al mercado: valor de mercado: el enigma del malentendido, *Journal of Property Investment and Finance*, vol. 39, núm. 5, 2021, págs. 492–499.

FRENCH N., GABRIELLI L., Precios de mercado: valoración de propiedades revisada: la jerarquía de enfoques, métodos y modelos de valoración, *Journal of Property Investment & Finance*, vol. 36, núm. 4, 2018, págs. 391–396.

GHOLIPOUR FH, TAJADDINI R., PHAM TNT, Transparencia del mercado inmobiliario e incumplimiento de las hipotecas *Investigación en Negocios y Finanzas Internacionales* 53 10120, 2020.

GHOLIPOUR FH, MASRON AT, Factores del mercado inmobiliario e inversión inmobiliaria extranjera. *J. Economía. Semental.* 40 (4), 2013, págs. 448–468.

GHOSH M., GUHA R., SARKAR R. Y ABRAHAM A., Una técnica de selección de características de filtro envolvente basada en la optimización de colonias de hormigas, *Neural Computing and Applications*, vol. 32, núm. 12, 2020, págs. 7839–7857.

GORDON BL Y WINKLER DT, El efecto de los cambios en los precios de cotización en el precio de venta de viviendas residenciales unifamiliares, *The Journal of Real Estate Finance and Economics*, 2016, págs. 1–31.

GUERRIERI G., L'informazione per l'efficienza e la trasparenza del mercato immobiliare: l'esperienza italiana, *Territorio Italia*, n. 1, 2011, págs. 88–102.

HAYUNGA DK Y PACE RK, Precios de lista en el mercado inmobiliario de EE. UU., *The Journal of Real Estate Finance and Economics*, 2016, págs. 1–30.

INTERNATIONAL VALUATION STARDARD COUNCIL IVSC, *International Valuation Standards*, Londres, 2020.

IONAS, CU E., ANGHEL I., Mejora de la transparencia inmobiliaria mediante la digitalización, *Actas de la Conferencia Internacional sobre Excelencia Empresarial* vol. 14(1), julio de 2020, págs. 371–384.

IONAS, CU E., TALTAVULL DE LA PAZ P. Y MIRONIUC M., La relación entre precios de la vivienda y transparencia del mercado. Evidencia de los mercados europeos metropolitanos, *Vivienda, teoría y sociedad*, vol. 38, núm. 1, 2021, págs. 42–71.

JOHN LANG LASALLE, Índice global de transparencia inmobiliaria, 2022 - Transparencia en una era de incertidumbre, Informe de transparencia inmobiliaria, disponible en: [www.joneslanglasalle.com](http://www.joneslanglasalle.com) (consultado en julio de 2022).

LINDQVIST S., El concepto de transparencia en el mercado inmobiliario residencial de la Unión Europea: un marco teórico, *Revista Internacional de Derecho en el Medio Ambiente Construido*, vol. 4, 2012, págs. 99–115.

MARELI M. Y TWALA B., Un algoritmo de búsqueda adaptativo de cuco para optimización, *Computación e informática aplicadas*, vol. 14, núm. 2, 2018, págs. 107–115.

NEWELL G., La cambiante transparencia del mercado inmobiliario en los mercados inmobiliarios europeos. *J. Prop. Invertir. Finanzas.* 34 (4), 2016, págs. 407–420.

PITTARELLO M., SCARPA M., RUGGERI AG, GABRIELLI L. Y

[

SCHIBUOLA L., Redes neuronales artificiales para optimizar proyectos de edificios de energía cero (ZEB) desde las primeras etapas de diseño, *Ciencias Aplicadas*, 11, 2021, p. 5377.

Pozo AG, Una estructura anidada del mercado inmobiliario: evidencia adicional, *Estudios de Vivienda*, vol. 24, núm. 3, 2009, págs. 373–395.

RAZALI MN Y ADNAN YM, Transparencia en las empresas inmobiliarias de Malasia, *Property Management*, 30(5), 2012, págs. 398–415.

SADAYUKI T., HARANO K. Y YAMAZAKI F., Transparencia del mercado e inversión inmobiliaria internacional, *Journal of Property Investment and Finance*, vol. 37, núm. 5, 2019, págs. 503–518.

SCHULTE K.-W., ROTTKE N. Y PITCHKE C., Transparencia en el mercado inmobiliario alemán, *Journal of Property Investment and Finance*, vol. 23, núm. 1, 2005, págs. 90–108.

SIHAM A., SARA S. Y ABDELLAH A., Selección de funciones basada en el aprendizaje automático para la calificación crediticia: una evaluación de filtros y métodos integrados, *Conferencia internacional sobre innovaciones en sistemas y aplicaciones inteligentes (INISTA)*, 2021, págs. 1–6 .

SIMONOTTI M., *Metodi Di Stima Immobiliare*, Flaccovio, Palermo, 2006.

SURESH SMS Y NARAYANAN A., Mejora de la precisión de la clasificación mediante la técnica de selección de funciones combinadas de filtro y envoltura, *Conferencia internacional IEEE sobre tecnologías eléctricas, informáticas y de comunicación*

(ICECCT), 2019, págs. 1–6.

TATWANI S. Y KUMAR E., Comparación paramétrica de varias técnicas de selección de características, *Journal of Advanced Research in Dynamical and Control Systems*, vol. 11, núm. 10 Número especial, 2019, págs. 1180–1190.

UGOLINI M., *Metodologie di apprendimento automatico applicate alla generazione di dati 3d*, 2014, disponible en <https://amslaurea.unibo.it/10415/>.

UNI 11612:2015, Determinación del valor de mercado de las propiedades, 2015.

UNI/PdR 53:2019, Análisis del mercado inmobiliario - Directrices para identificar el segmento de mercado y recopilar datos inmobiliarios, 2019.

WANG A. AND XU Y., Análisis de regresión lineal múltiple del precio inmobiliario, en IEEE (Ed.), *Conferencia Internacional sobre Robots y Sistemas Inteligentes, ICRIS 2018, Changsha (China)*, 2018, págs.

YASSI M. Y MOATTAR MH, Selección de características robusta y estable mediante la integración de métodos de clasificación y técnica de envoltura en la clasificación de datos genéticos, *Comunicaciones de investigación bioquímica y biofísica*, vol. 446, núm. 4, 2014, págs. 850–856.

YUN L. Y CHAU KW, El impacto de la transparencia del mercado inmobiliario en los vínculos entre los bienes raíces indirectos y directos, documento presentado en la Conferencia ERES, Viena, julio de 2013, págs. 3–6.

[