



Facultad de Ciencias Económicas y Empresariales

# **APLICACIÓN DE *MACHINE* *LEARNING* PARA LA DETECCIÓN DE BURBUJAS INMOBILIARIAS**

Clave: 202005834

## **RESUMEN**

La volatilidad en los precios de la vivienda y su impacto en la estabilidad financiera han generado un creciente interés en el desarrollo de herramientas predictivas que permitan anticipar episodios de sobrevaloración y prevenir crisis en la economía de los países. En este contexto, el presente trabajo analiza la aplicación de Modelos de Cambio de Régimen de Markov y algoritmos de Machine Learning para la detección temprana de burbujas inmobiliarias en el mercado de Estados Unidos.

A través del uso de indicadores macroeconómicos clave y técnicas de aprendizaje supervisado, se ha desarrollado un marco metodológico que permite identificar patrones de sobrevaloración y evaluar el riesgo de correcciones en los precios de la vivienda. Estos mecanismos de alerta temprana pueden aplicarse en distintos ámbitos. Para reguladores financieros, representan herramientas clave en la evaluación de riesgos y la formulación de políticas macroprudenciales. Para inversores institucionales y gestores de activos, ofrecen la posibilidad de optimizar estrategias de inversión y mitigar los riesgos asociados a la exposición a mercados sobrevalorados. Asimismo, su integración en la planificación urbana y los mercados hipotecarios podría contribuir a un desarrollo más equilibrado del sector inmobiliario, fortaleciendo la estabilidad económica a largo plazo.

## **PALABRAS CLAVE**

Mercado inmobiliario, burbuja especulativa, modelo de cambio de régimen, Machine Learning, alerta temprana

## **ABSTRACT**

Housing price volatility and its impact on financial stability have sparked growing interest in developing predictive tools that can anticipate overvaluation episodes and prevent recessions in countries' economies. In this context, this study examines the application of Markov Regime-Switching Models and Machine Learning algorithms for the early detection of real estate bubbles in the United States' market.

By leveraging key macroeconomic indicators and supervised learning techniques, a methodological framework has been developed to identify overvaluation patterns and assess the risk of housing price corrections. These early warning mechanisms can be applied across various domains. For financial regulators, they serve as essential tools in risk assessment and macroprudential policy formulation. For institutional investors and asset managers, they provide the opportunity to optimize investment strategies and mitigate the risks associated with the exposure to overvalued markets. Furthermore, their integration into urban planning and the mortgage markets could contribute to a more balanced development of the real estate sector, reinforcing long-term economic stability.

## **KEY WORDS**

Real estate market, speculative bubble, regime-switching model, Machine Learning, early warning

## ÍNDICE DE CONTENIDO

<b>I. INTRODUCCIÓN</b> .....	7
1. PLANTEAMIENTO Y JUSTIFICACIÓN DE LA CUESTIÓN.....	7
2. OBJETIVOS DE LA INVESTIGACIÓN.....	7
<b>II. MARCO TEÓRICO</b> .....	8
1. CONCEPTO DE BURBUJA INMOBILIARIA.....	8
2. DETERMINANTES ECONÓMICOS Y FINANCIEROS.....	9
<b>III. ESTADO DEL ARTE</b> .....	10
<b>IV. METODOLOGÍA</b> .....	12
1. EXTRACCIÓN DE DATOS Y SELECCIÓN DE VARIABLES.....	12
1.1. <i>Case-Shiller Index</i> .....	13
1.2. <b>Variables explicativas</b> .....	14
2. PREPROCESAMIENTO DE DATOS.....	15
2.1. <b>Deflactación de variables económicas</b> .....	15
2.2. <b>Conversión de variables cuatrimestrales a frecuencia mensual</b> .....	17
3. ANÁLISIS EXPLORATORIO-DESCRIPTIVO.....	17
3.1. <b>Estadísticos principales, identificación de valores atípicos y evolución temporal de las variables</b> .....	18
3.2. <b>Correlación entre variables</b> .....	20
4. MODELO DE CAMBIO DE RÉGIMEN DE MARKOV.....	21
4.1. <b>Introducción</b> .....	21
4.2. <b>Diseño e implementación</b> .....	22
5. MODELOS DE MACHINE LEARNING.....	23
5.1. <b>Algoritmos empleados</b> .....	23
5.1.1. <i>Random Forest</i> .....	23
5.1.2. <i>XGBoost (Extreme Gradient Boosting)</i> .....	24
5.1.3. <i>Redes neuronales artificiales</i> .....	25

5.2.	<b>Diseño e implementación</b> .....	26
V.	<b>ANÁLISIS DE RESULTADOS EMPÍRICOS</b> .....	26
1.	MODELO DE CAMBIO DE RÉGIMEN DE MARKOV.....	26
1.1.	<b>Principios de los 90</b> .....	28
1.2.	<b>Crisis inmobiliaria de las hipotecas <i>suprime</i></b> .....	29
1.3.	<b>Post-crisis y recuperación gradual</b> .....	30
1.4.	<b>Impacto de la pandemia de COVID-19</b> .....	31
1.5.	<b>Tendencias actuales</b> .....	32
2.	EVALUACIÓN Y COMPARACIÓN DE LOS ALGORITMOS DE <i>MACHINE LEARNING</i> .....	32
2.1.	<b>Descripción de métricas de evaluación empleadas</b> .....	33
2.1.1.	<i>Error absoluto medio (Mean Absolute Error, MAE)</i> .....	33
2.1.2.	<i>Error cuadrático medio (Mean Squared Error, MSE) y raíz del error cuadrático medio (Root Mean Squared Error, RMSE)</i> .....	33
2.1.3.	<i>Error cuadrático medio ponderado (Weighted Mean Squared Error, WMSE)</i> .....	34
2.1.4.	<i>Coefficiente de determinación (<math>R^2</math>)</i> .....	34
2.1.5.	<i>Correlación de Pearson</i> .....	34
2.2.	<b>Análisis comparativo del desempeño de los algoritmos</b> .....	34
VI.	<b>CONCLUSIONES</b> .....	37
1.	PROPUESTAS DE APLICACIÓN.....	38
2.	POTENCIALES LIMITACIONES.....	38
VII.	<b>DECLARACIÓN DEL USO DE HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL GENERATIVA</b> .....	41
VIII.	<b>BIBLIOGRAFÍA</b> .....	43
IX.	<b>ANEXOS</b> .....	47
1.	BASE DE DATOS DE EXCEL .....	47
2.	CÓDIGO DE PYTHON.....	47

## ÍNDICE DE IMÁGENES

Imagen 1: Modelo de cambio de régimen de Markov con (arriba) vs. Sin (abajo) deflactación del Case-Shiller Index .....	16
Imagen 2: Comparación frecuencia trimestral vs. mensual de la variable MORTG_DEBT .....	17
Imagen 3: Estadísticos principales de las variables explicativas.....	18
Imagen 4: Boxplots de las variables explicativas.....	18
Imagen 5: Gráficos de evolución temporal de las variables explicativas.....	19
Imagen 6:Matriz de correlación de las variables explicativas.....	20
Imagen 7: Esquema del funcionamiento del algoritmo Random Forest .....	24
Imagen 8: Esquema de funcionamiento del algoritmo ANN.....	25
Imagen 9: Identificación de ciclos Inmobiliarios según Markov Switching Model.....	27
Imagen 10: Probabilidad de régimen de colapso de burbuja (crash) según Markov Switching Model .....	27
Imagen 11: Comparación de métricas de evaluación de los modelos de Machine Learning .....	35
Imagen 12: Scatter plot comparativo de las predicciones de los modelos de Machine..	36

## ÍNDICE DE TABLAS

Tabla 1: Descripción de las variables explicativas empleadas .....	13
Tabla 2: Descripción estados modelo de cambio de régimen de Markov .....	22
Tabla 3: Resumen métricas de evaluación de los modelos de Machine Learning .....	35

# I. INTRODUCCIÓN

## 1. PLANTEAMIENTO Y JUSTIFICACIÓN DE LA CUESTIÓN

El mercado inmobiliario es un pilar fundamental en la economía de cualquier país, dado su impacto en el crecimiento económico, la estabilidad financiera y el bienestar social. Sin embargo, es también un sector vulnerable a la formación de burbujas especulativas, caracterizadas por un aumento excesivo y sostenido en los precios de los activos inmobiliarios, seguido de una corrección abrupta.

A lo largo de la historia, las burbujas inmobiliarias han tenido consecuencias devastadoras. Evidencia de ello es la crisis financiera de 2008, cuyo catalizador principal fue el colapso del mercado hipotecario en Estados Unidos. Estos episodios muestran la necesidad de desarrollar herramientas analíticas que permitan identificar patrones de comportamiento en los mercados y anticipar posibles escenarios de sobrevaloración con el fin de mejorar la toma de decisiones preventivas.

Así, el presente trabajo aborda la temática relativa a la identificación y predicción de burbujas inmobiliarias mediante el uso de modelos de *Machine Learning*. A través de un doble enfoque teórico y empírico, se busca evaluar la eficacia de distintas metodologías para detectar fases de exuberancia en los mercados inmobiliarios y anticipar posibles crisis. La detección temprana puede desempeñar un papel clave en la mitigación de estos fenómenos, contribuyendo a desinflar las burbujas antes de que colapsen. De esta manera, se pretende contribuir a la estabilidad del sector y proporcionar información valiosa para la toma de decisiones informadas por parte de reguladores e inversores.

## 2. OBJETIVOS DE LA INVESTIGACIÓN

El objetivo principal de este trabajo de investigación es analizar la formación de burbujas inmobiliarias y evaluar la capacidad de distintos modelos de *Machine Learning* para predecirlas. Para ello, en primer lugar, se busca delimitar con precisión el concepto de burbuja inmobiliaria y sus principales características, diferenciándolo de ciclos normales de crecimiento del mercado e identificando los principales factores económicos y financieros que influyen en su formación.

Para ello, se procede a la recopilación y análisis de datos relevantes, explorando la relación entre distintas variables macroeconómicas. A partir de esta base de datos, se lleva

a cabo la implementación de distintos enfoques metodológicos de *Machine Learning*, evaluando su desempeño predictivo en la identificación de fases de exuberancia en los mercados inmobiliarios.

Finalmente, a partir de los resultados obtenidos, se proponen estrategias basadas en las herramientas desarrolladas que puedan ser utilizadas tanto por reguladores del mercado inmobiliario como por inversores para mejorar la capacidad de anticipación y gestión de los riesgos asociados a la formación de burbujas en el sector inmobiliario.

## II. MARCO TEÓRICO

### 1. CONCEPTO DE BURBUJA INMOBILIARIA

Como primer paso esencial para este estudio, es fundamental delimitar con precisión el concepto de burbuja inmobiliaria. Aunque se emplea con frecuencia, su significado no siempre es claro. Por ello, en este apartado se analiza el concepto desde un enfoque teórico, examinando los principales determinantes económicos y financieros identificados en la literatura que explican su formación y desarrollo.

El estudio de las burbujas inmobiliarias es intrínsecamente complejo debido a la doble naturaleza de la vivienda. Por un lado, actúa como un bien de consumo que satisface necesidades habitacionales y, por otro, desempeña un papel como activo financiero con potencial de revalorización y reventa. Esta dualidad da lugar a un mercado en el que participan tanto propietarios cuyo interés principal es el uso y disfrute de la vivienda como especuladores que adquieren propiedades con fines especulativos (Álvarez-Lois y Nuño-Barrau, 2007).

En un escenario de equilibrio, el precio de la vivienda debería reflejar únicamente sus fundamentos económicos a largo plazo, de modo que cualquier variación en su valor esté vinculada a cambios en tales determinantes. No obstante, en la práctica, el mercado inmobiliario muestra que los precios pueden desviarse temporalmente de su nivel de equilibrio debido a restricciones en la oferta que dificultan una respuesta inmediata a un incremento en la demanda. En estos casos, se genera una sobrevaloración transitoria, seguida de un ajuste progresivo a medida que el mercado se estabiliza (Ayuso y Restoy, 2006). En este sentido, es importante recalcar que, a diferencia de otros tipos de activos, el mercado inmobiliario se distingue por una oferta inelástica en el corto plazo,

Por otro lado, el precio de la vivienda también puede aumentar impulsado por expectativas de incrementos futuros, lo que fomenta la demanda especulativa y genera un ciclo en el que los compradores adquieren inmuebles con la esperanza de venderlos a un precio aún mayor. Este fenómeno provoca una escalada de precios que no sólo supera el nivel de equilibrio del mercado, sino que también excede la trayectoria de ajuste natural de una sobrevaloración. En este contexto, surge una burbuja inmobiliaria propiamente dicha, ya que el mercado se vuelve vulnerable a un colapso repentino si las expectativas cambian abruptamente (Ayuso y Restoy, 2006).

En definitiva, una burbuja inmobiliaria se produce cuando los precios de la vivienda se alejan significativamente de su valor intrínseco, impulsados más por expectativas de revalorización que por fundamentos económicos sólidos (Martínez García, 2024). Este fenómeno distorsiona el mercado, lo que eventualmente conduce al colapso de la burbuja cuando los precios dejan de sostenerse y experimentan una corrección repentina.

## 2. DETERMINANTES ECONÓMICOS Y FINANCIEROS

Dada la definición de burbuja inmobiliaria propuesta, resulta fundamental comprender los factores económicos y financieros que explican el valor intrínseco de la vivienda y, en consecuencia, permiten identificar cuándo su precio se desvía de estos fundamentos, dando lugar a una burbuja. En este sentido, es abundante la literatura que ha abordado esta cuestión y, de manera general, clasifica los determinantes económicos del mercado inmobiliario en dos grandes grupos: aquellos que afectan a la demanda y aquellos que inciden sobre la oferta.

Desde la perspectiva de la demanda, el precio de la vivienda está influenciado principalmente por la renta disponible de los hogares, los tipos de interés y ciertas variables demográficas como la población o la tasa de desempleo (Álvarez-Román y García-Posada, 2021).

En este sentido, un aumento en el volumen de ingresos de la población o una disminución en la tasa de desempleo tienden a elevar la demanda de vivienda, impulsando los precios al alza. Del mismo modo, tipos de interés más bajos reducen el coste de acceder a financiamiento hipotecario, incentivando la compra de inmuebles. En este contexto, diversos autores destacan la expansión del crédito y el consiguiente nivel de endeudamiento de los hogares como una variable clave, ya que una proporción elevada

de deuda en relación con los ingresos puede alimentar ciclos especulativos y aumentar la vulnerabilidad del mercado ante cambios en las condiciones económicas de los hogares (Martínez García, 2024).

Por otro lado, aunque a menudo recibe menos atención, el papel de la oferta es también clave en la dinámica del mercado inmobiliario. Como se mencionó anteriormente, la oferta de vivienda es altamente inelástica en el corto plazo, lo que significa que no puede ajustarse rápidamente a variaciones en la demanda. Esta rigidez amplifica los ciclos de auge y caída, permitiendo que los precios aumenten bruscamente cuando la demanda crece, pero retrasando la corrección cuando las condiciones cambian. En este sentido, el stock de vivienda, las restricciones urbanísticas y los plazos administrativos en los procesos de construcción y desarrollo son factores esenciales que inciden en la posible aparición de burbujas inmobiliarias (García-Montalvo, 2006).

En otro orden de cosas, algunos autores identifican la volatilidad del mercado bursátil como otro factor influyente en la exuberancia del mercado inmobiliario, ya que puede llevar a los inversores a refugiarse en bienes raíces en busca de estabilidad o mayores rendimientos, lo que contribuye al alza de los precios. En períodos de incertidumbre y volatilidad en los mercados, la vivienda se percibe como un activo seguro, atrayendo más inversión y sosteniendo una escalada de precios incluso sin un respaldo en sus fundamentos económicos (García-Montalvo, 2006).

### **III. ESTADO DEL ARTE**

Como paso previo a la elección de la metodología empleada en este estudio, se ha llevado a cabo una revisión de la literatura relevante con el objetivo de identificar las principales técnicas utilizadas por los economistas para la predicción de burbujas inmobiliarias especulativas.

García-Montalvo (2006) identifica dos enfoques principales para la detección de burbujas inmobiliarias. El primero se basa en el uso de modelos econométricos para separar el valor fundamental del activo, determinado por los factores económicos expuestos, del componente especulativo o burbuja. Este método consiste en estimar un precio "fundamental" de la vivienda y calcular sus desviaciones respecto a los precios observados, siendo una estrategia ampliamente utilizada por los bancos centrales (Cañizares Martínez, 2023).

El segundo enfoque se centra en la extracción de expectativas de revalorización de los compradores mediante encuestas, bajo la premisa de que las burbujas surgen, en gran medida, por expectativas de apreciación futura. En esta línea, en su estudio de 2006, en plena fase de expansión del mercado inmobiliario español, García-Montalvo realizó una encuesta a 1.509 españoles para analizar los motivos detrás de la compra de vivienda y la importancia de la rentabilidad esperada en la decisión de adquisición. Los resultados mostraron que el 65% de los encuestados consideraba la elevada rentabilidad de los activos inmobiliarios como un factor clave o, al menos, relevante en su decisión de compra, lo que evidencia el papel de las expectativas en la demanda de vivienda (García-Montalvo, 2006).

No obstante, uno de los principales desafíos en la detección de burbujas inmobiliarias ha sido la dificultad de modelar con precisión el valor intrínseco de los activos, lo que ha conducido a estimaciones sesgadas y un reconocimiento tardío *ex post* de estos fenómenos. En respuesta, los avances recientes en técnicas de series de tiempo y paneles de datos han permitido el desarrollo de métodos más precisos para identificar burbujas en tiempo real.

En este sentido, Cañizares Martínez (2023) identifica dos enfoques principales. El primero se basa en la identificación de comportamientos explosivos en los precios, siguiendo el método propuesto por Phillips et al. (2011, 2015), que emplea pruebas estadísticas avanzadas (como versiones ampliadas del test de Dickey-Fuller) para detectar episodios de exuberancia en los precios de los activos. El segundo enfoque es el empleo de modelos de cambio de régimen (*regime-switching models*), introducidos por Van Norden y Schaller (1993, 1996), que permiten detectar períodos de alta y baja valoración en los precios de la vivienda, identificando patrones cíclicos en el mercado inmobiliario (Cañizares Martínez, 2023). Estas técnicas se centran en patrones estadísticos indicativos de burbujas sin necesidad de modelar explícitamente el valor intrínseco de los activos, lo que reduce el riesgo de errores de estimación y permite una detección más temprana de episodios de sobrevaloración.

En esta línea, Cañizares Martínez (2023) desarrolla un modelo de cambio de régimen de Markov para analizar la evolución del mercado inmobiliario estadounidense en función del crédito hipotecario. Su enfoque resalta el papel de la deuda como un factor clave en la dinámica de los precios de la vivienda, facilitando la identificación de transiciones entre distintos regímenes, como estabilidad, auge y crisis.

Por su parte, Fabozzi et al. (2020) aplican un modelo híbrido para la detección de burbujas inmobiliarias en EE.UU. y Reino Unido, combinando pruebas de raíces unitarias para identificar comportamientos explosivos con un modelo de cambio de régimen. Su metodología segmenta el mercado en dos estados principales: uno en el que la burbuja se mantiene y sigue creciendo, y otro en el que colapsa.

En otro orden de cosas, en los últimos años, la aplicación de técnicas de *Machine Learning* en el sector inmobiliario ha avanzado significativamente. Ejemplo de ello es el estudio de Zhang (2021), que construye un algoritmo de regresión lineal múltiple para predecir el precio de la vivienda a partir de variables como la superficie, el número de habitaciones y el índice de criminalidad de la ciudad, basándose en un conjunto de datos de precios de vivienda en Boston.

Por su parte, Bokreta y Leblanc (2009) centran su investigación en la predicción del *Case-Shiller Index* empleando técnicas de regresión (específicamente los métodos de *Ordinary Least Squares* (OLS) y *Kalman filte*) con el objetivo de mejorar la precisión en la estimación de la evolución del mercado inmobiliario.

En resumen, la literatura ha desarrollado múltiples enfoques para la detección de burbujas inmobiliarias, desde modelos econométricos tradicionales hasta técnicas avanzadas de series temporales, cambio de régimen y aprendizaje automático.

#### **IV. METODOLOGÍA**

Este apartado describe el proceso metodológico seguido en este estudio, desde la recopilación y preprocesamiento de datos hasta la implementación de los modelos predictivos.

##### **1. EXTRACCIÓN DE DATOS Y SELECCIÓN DE VARIABLES**

Los datos empleados han sido obtenidos del portal online *Federal Reserve Economic Data* (FRED), gestionado por el Departamento de Investigación de la Reserva Federal de St. Louis. Se han descargado en archivos .xls los datos correspondientes al período comprendido entre el 1 de enero de 1987 (fecha correspondiente al primer dato disponible del *Case-Shiller Index*) y el 1 de julio de 2024, abarcando un total de 451 observaciones.

Variable	Acrónimo	Tipo	Unidades	Frecuencia
<a href="#">S&amp;P CoreLogic Case-Shiller U.S. National Home Price Index</a>	CS_INDEX	Numérica	Index Jan 2000=100	Mensual
<a href="#">Consumer Price Index</a>	CPI	Numérica	Index 1982-1984=100	Mensual
<a href="#">Tasa de desempleo</a>	UNRATE	Numérica	%	Mensual
<a href="#">Sueldos y salarios</a>	WAGES_SAL	Numérica	Billones de dólares	Mensual
<a href="#">Tipo de interés hipotecario a 30 años</a>	MORTGAGE30	Numérica	%	Mensual
<a href="#">Tasa efectiva de fondos federales de EEUU</a>	FEDFUNDS	Numérica	%	Mensual
<a href="#">Índice de Producción Industrial</a>	INDPRO	Numérica	Index 2017=100	Mensual
<a href="#">Población en edad de trabajar (15-64 años)</a>	WORK_AGE_POP	Numérica	Número de personas	Mensual
<a href="#">Precio de alquiler medio de la residencia principal</a>	RENT	Numérica	Index 1982-1984=100	Mensual
<a href="#">Deuda hipotecaria</a>	MORTG_DEBT	Numérica	Millones de dólares	Trimestral
<a href="#">Nuevas viviendas privadas en construcción</a>	UND_CONSTR	Numérica	Número de viviendas	Mensual

*Tabla 1: Descripción de las variables explicativas empleadas*

*Fuente: Elaboración propia*

### **1.1. Case-Shiller Index**

El principal indicador utilizado en la metodología propuesta es el Índice Nacional de Precios de Vivienda S&P CoreLogic Case-Shiller de EE.UU (*Case-Shiller Index*), un

barómetro ampliamente reconocido del comportamiento del mercado inmobiliario estadounidense y considerado actualmente como el más preciso para evaluar la evolución de los precios de la vivienda a lo largo del tiempo (Coleman et al., 2008).

Este índice fue desarrollado por los economistas Karl E. Case y Robert J. Shiller y es publicado mensualmente por la agencia S&P Dow Jones Indices. Su objetivo es medir la evolución del valor de las viviendas unifamiliares en las 9 divisiones del censo de EE.UU. mediante un algoritmo de media móvil de 3 meses. De este modo, los cambios porcentuales en el índice reflejan variaciones en los precios del mercado de la vivienda, manteniendo constante la calidad de los inmuebles analizados. Para garantizar la precisión del indicador, se excluyen específicamente de los cálculos los cambios en los tipos, tamaños y características físicas de las viviendas, evitando así distorsiones en la medición del valor real del mercado inmobiliario (S&P Dow Jones Indices, 2024).

El *Case-Shiller Index* se compone de tres medidas principales:

- **Índice Nacional:** cubre el mercado de la vivienda en todo el país.
- **Índice de 10 ciudades (*10-City Composite Index*):** analiza el comportamiento de los precios en las principales áreas metropolitanas.
- **Índice de 20 ciudades (*20-City Composite Index*):** amplía la cobertura a un mayor número de mercados urbanos clave.

Para este estudio se ha empleado como indicador el Índice Nacional, ya que proporciona una visión integral y representativa de la evolución del mercado inmobiliario en EE.UU.

## 1.2. Variables explicativas

Para la selección de las variables explicativas utilizadas en los modelos predictivos de *Machine Learning* se ha llevado a cabo un análisis exhaustivo de los indicadores macroeconómicos empleados en estudios previos, destacando las investigaciones de Carlos Cañizares (2023); Álvarez-Lois y Nuño-Barrau (2007); Álvarez-Román y García-Posada (2021); y Nneji, Brooks y Ward (2013).

En base a ello, las variables explicativas seleccionadas han sido:

- **Tasa de desempleo:** indicador clave del mercado laboral, cuya evolución influye directamente en la capacidad de compra de los hogares.

- **Sueldos y salarios:** refleja el poder adquisitivo de la población, determinado la capacidad de los individuos para adquirir una vivienda.
- **Tasa de interés hipotecario a 30 años:** determina el coste de la financiación para la compra de vivienda.
- **Tasa efectiva de fondos federales:** incide en el coste general del crédito y, por ende, en las condiciones de financiamiento hipotecario.
- **Índice de Producción Industrial:** indicador de la actividad económica general, con impacto en el nivel de empleo y en la demanda de vivienda.
- **Población en edad de trabajar (15-64 años):** factor demográfico determinante, en cuanto una mayor población activa suele estar asociada a una mayor necesidad de inmuebles.
- **Precio de alquiler medio de la residencia principal:** variable clave para analizar la relación entre el mercado de compra y alquiler. Dado que la vivienda es un activo en la cartera de las familias que genera rendimientos (ya sea por alquiler o por uso habitacional), en el equilibrio el precio del alquiler debería reflejar el valor de estos servicios, en cuanto que los individuos pueden elegir entre comprar o alquilar para obtener el mismo beneficio (Ayuso y Restoy, 2006).
- **Deuda hipotecaria:** refleja la evolución del crédito inmobiliario, indicando el nivel de endeudamiento de los hogares y la disposición de las instituciones financieras a conceder préstamos.
- **Nuevas viviendas privadas en construcción:** indicador clave del dinamismo de la oferta del sector inmobiliario y su capacidad para responder a la demanda de vivienda.

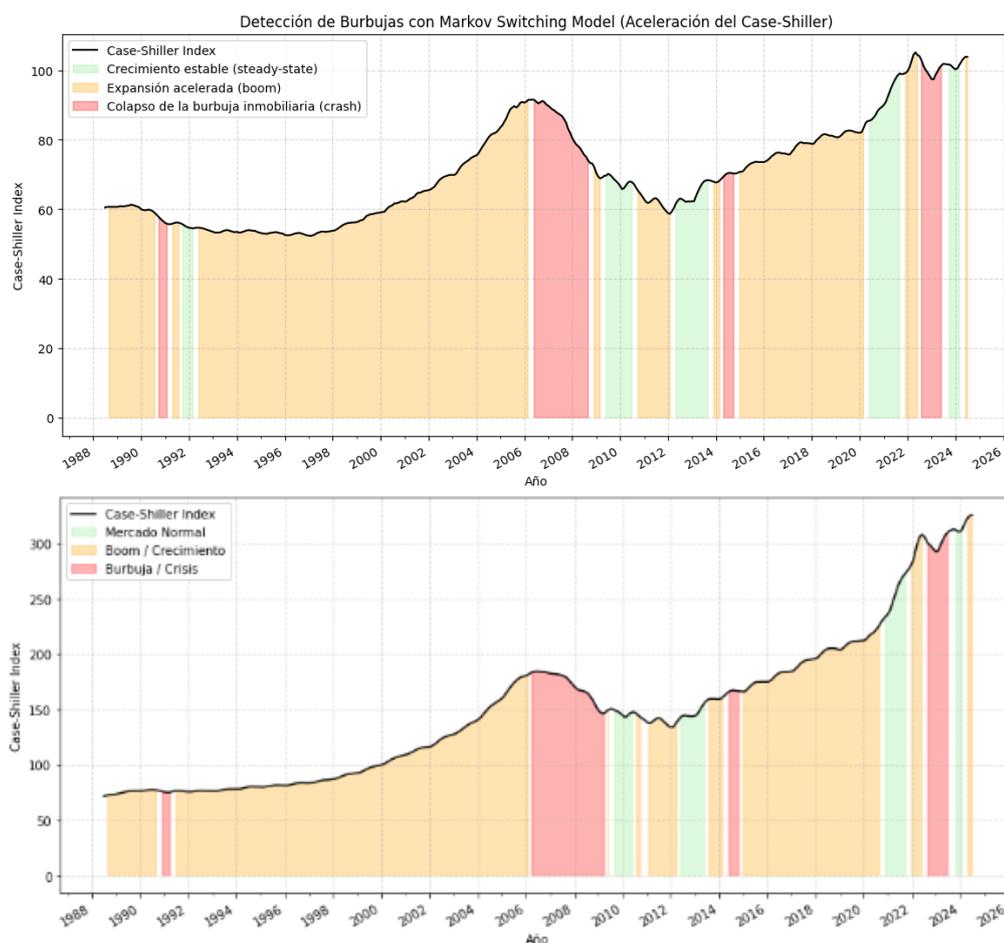
Asimismo, si bien no se ha empleado como una variable explicativa *per se* en los modelos de *Machine Learning*, también se ha procedido a la recolección del Índice de Precios del Consumidor (*Consumer Price Index*, CPI) con el objetivo de ajustar ciertas variables en base a la inflación, lo cual será detallado en apartados posteriores.

## 2. PREPROCESAMIENTO DE DATOS

### 2.1. Deflatación de variables económicas

Existe un debate académico sobre la conveniencia de ajustar los precios de la vivienda en base a la inflación. Algunos autores sostienen que este ajuste es fundamental para

diferenciar los aumentos reales en el valor de los inmuebles de los efectos generales del crecimiento del nivel de precios, mientras que otros argumentan que el análisis en términos nominales es suficiente para capturar la percepción de compradores y vendedores en el mercado. Ante esta dualidad de enfoques, este estudio ha considerado ambas perspectivas y ha optado por evaluar los resultados con y sin ajuste inflacionario. Para ello, se ha realizado la transformación del *Case-Shiller Index* a términos reales mediante la deflatación en base al CPI, siguiendo el enfoque propuesto por Carlos Cañizares (2023, p. 23), y Álvarez-Román y García-Posada (2021, p.4) en sus estudios.



*Imagen 1: Modelo de cambio de régimen de Markov con (arriba) vs. Sin (abajo) deflatación del Case-Shiller Index*

*Fuente: Elaboración propia en Python*

Como se muestra en la Imagen 1, los resultados del Modelo de Cambio de Régimen de Markov son similares tanto con el *Case-Shiller Index* sin deflactar como ajustado por la inflación. Dado que la mayoría de la literatura destaca la importancia de monitorear los precios reales para detectar burbujas inmobiliarias, este estudio adopta el segundo

enfoque. Como señala Martínez García (2024), un aumento excesivo en los precios nominales puede responder a un repunte inflacionario en lugar de una burbuja, por lo que expresarlos en términos reales ayuda a evitar confusiones en entornos de alta inflación (Martínez García, 2024).

Asimismo, esta deflatación se ha aplicado, no sólo al *Case-Shiller Index* en el modelo de Markov, sino también a las siguientes variables explicativas utilizadas en los algoritmos de Machine Learning: sueldos y salarios, precio del alquiler medio de la residencia principal y deuda hipotecaria.

## 2.2. Conversión de variables cuatrimestrales a frecuencia mensual

La variable “deuda hipotecaria” solo está disponible en el portal online FRED con una frecuencia trimestral. No obstante, para realizar un análisis más detallado y con un mayor número de observaciones, se ha decidido convertir estos datos a una frecuencia mensual mediante una interpolación lineal. Este enfoque sigue la metodología propuesta y aplicada por Carlos Cañizares (2023). Si bien esta conversión puede implicar ciertas limitaciones, dada la relevancia de la variable para la investigación, se ha optado por su implementación con el fin de obtener una mayor granularidad en los datos.

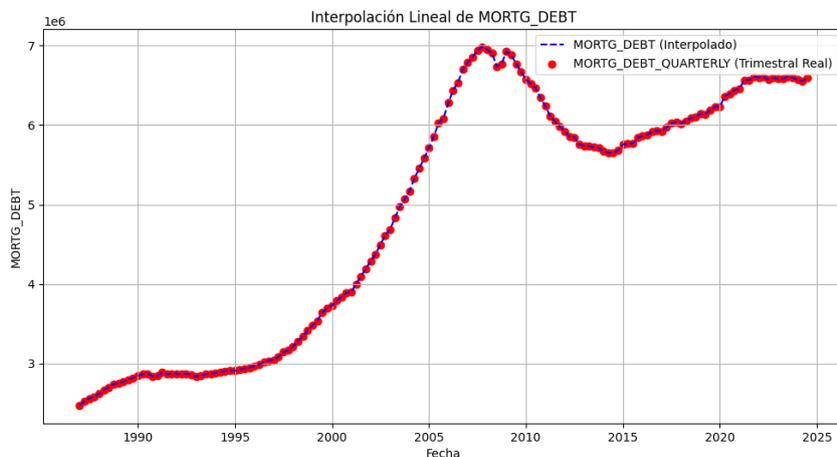


Imagen 2: Comparación frecuencia trimestral vs. mensual de la variable *MORTG\_DEBT*

Fuente: Elaboración propia en Python

## 3. ANÁLISIS EXPLORATORIO-DESCRIPTIVO

Una vez preprocesados los datos, se ha llevado a cabo un análisis exploratorio-descriptivo del dataset con el objetivo de comprender la distribución de las variables,

detectar posibles valores atípicos y evaluar las relaciones entre las variables explicativas y el *Case-Shiller Index*.

### 3.1. Estadísticos principales, identificación de valores atípicos y evolución temporal de las variables

En primer lugar, se han generado estadísticas descriptivas, gráficos *box plot* para detectar *outliers* y gráficos de evolución temporal para analizar tendencias y variaciones en las 10 variables explicativas.

	CS_INDEX	UNRATE	WAGES_SAL	MORTGAGE30	FEDFUNDS	INDPRO	WORK_AGE_POP	RENT	UND_CONSTR	MORTG_DEBT
count	451.00	451.00	451.00	451.00	451.00	451.00	451.00	451.00	451.00	451.00
mean	70.30	5.73	2867.42	6.34	3.22	87.61	186647240.58	113.88	972.37	4890445.14
std	14.64	1.68	567.54	2.14	2.66	14.92	18059247.74	9.01	306.52	1566444.81
min	52.30	3.40	1952.50	2.70	0.10	56.00	153835800.00	103.00	414.00	2474569.60
25%	58.45	4.50	2299.95	4.35	0.40	75.10	168693800.00	106.50	765.00	3001604.05
50%	67.70	5.40	2899.70	6.30	3.00	92.70	191700500.00	111.80	980.00	5681596.40
75%	80.85	6.55	3293.05	7.80	5.30	99.85	203961150.00	119.95	1135.00	6276841.88
max	105.20	14.80	3960.30	11.30	9.90	104.10	209091200.00	134.30	1711.00	6985564.00

Imagen 3: Estadísticos principales de las variables explicativas

Fuente: Elaboración propia en Python

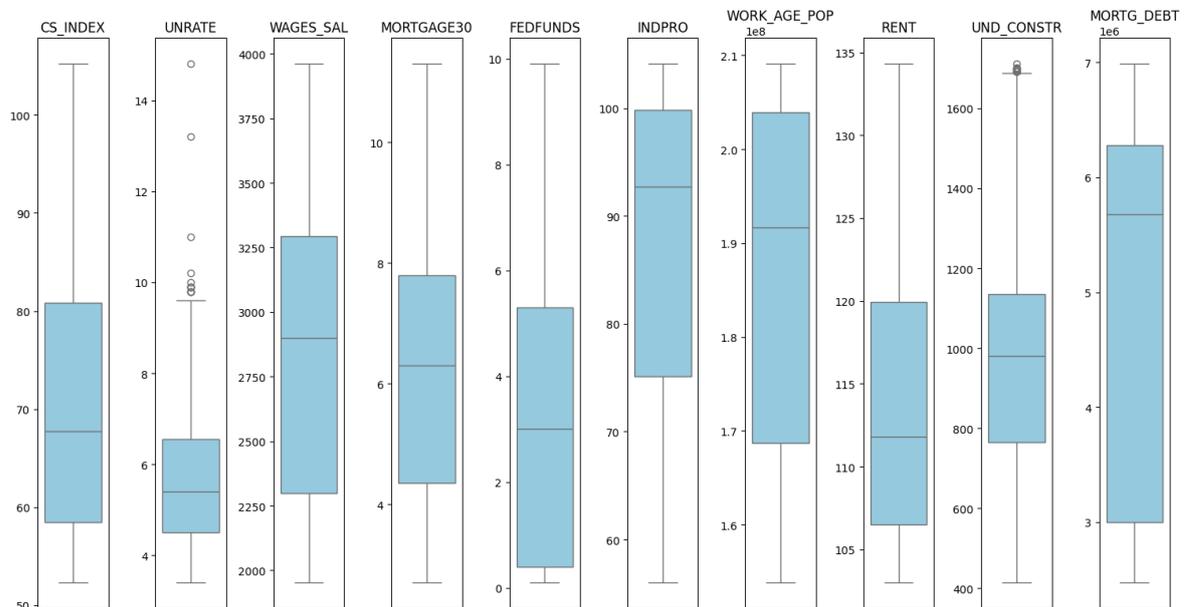


Imagen 4: Boxplots de las variables explicativas

Fuente: Elaboración propia en Python

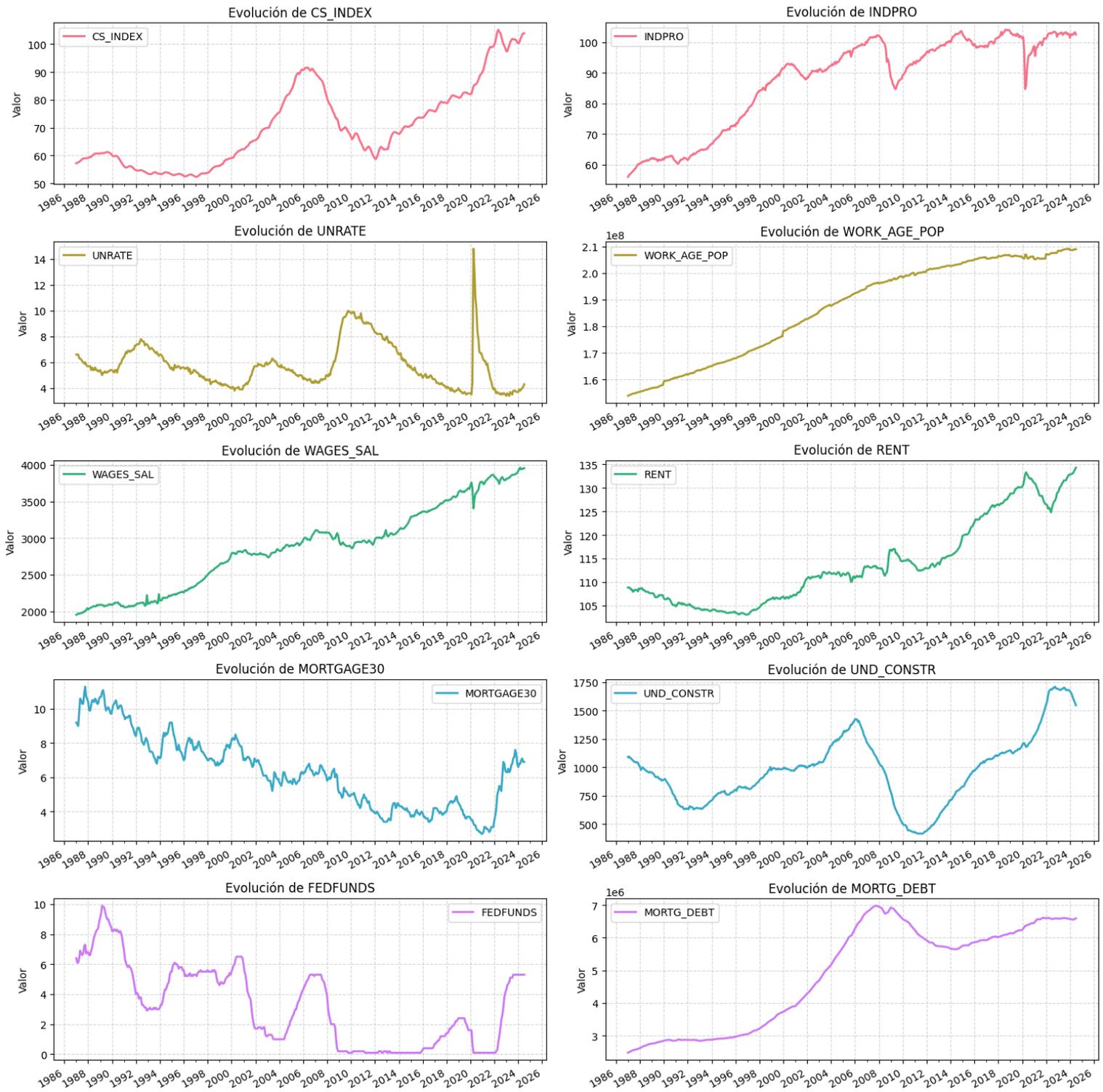


Imagen 5: Gráficos de evolución temporal de las variables explicativas

Fuente: Elaboración propia en Python

De acuerdo con las Imágenes 3 y 4, se han identificado posibles valores atípicos en las variables UNRATE (tasa de desempleo) y UND\_CONSTR (nuevas viviendas privadas en construcción).

En el caso de *UNRATE*, el valor máximo registrado (14.80%) supera significativamente el tercer cuartil, sugiriendo un posible *outlier*. Sin embargo, al verificar su origen, se confirmó que corresponde a abril de 2020, cuando la tasa de desempleo en EE.UU. alcanzó este nivel debido a la pandemia de COVID-19 según fuentes oficiales (U.S. Bureau of Labor Statistics, 2021). Dado que refleja un fenómeno económico real y documentado, no se considera necesario excluirlo del análisis. Por otro lado, el gráfico de la Imagen 5 muestra un crecimiento sostenido de *UND\_CONSTR* entre 2021 y 2023. Aunque se registran incrementos notables en la construcción de viviendas, parecen responder a una tendencia del mercado y no a anomalías estadísticas, por lo que no se considera necesario eliminar ningún *outlier* tampoco en este caso.

### 3.2. Correlación entre variables

Se ha realizado un análisis de correlación entre las variables explicativas y el *Case-Shiller Index* para evaluar la coherencia económica de sus relaciones.

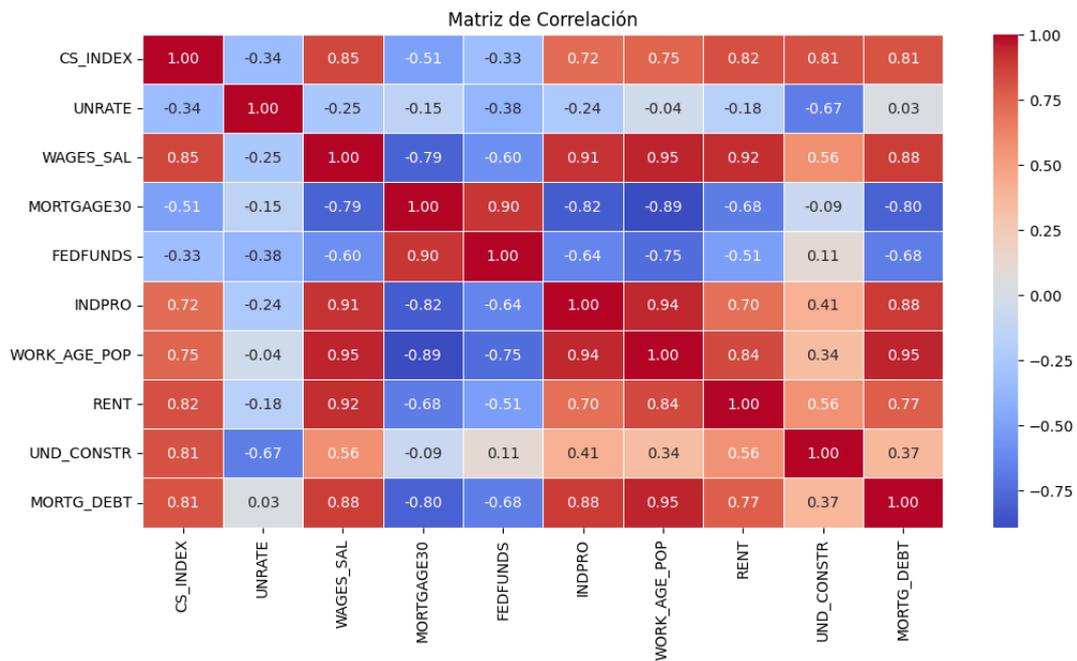


Imagen 6: Matriz de correlación de las variables explicativas

Fuente: Elaboración propia en Python

De acuerdo con la Imagen 6:

- **Tipos de interés:** el *Case-Shiller Index* muestra una correlación negativa con la tasa hipotecaria a 30 años y la tasa efectiva de fondos federales, lo que confirma que un

descenso en las tasas de interés abarata el crédito, aumenta la demanda de vivienda y presiona los precios al alza (Nneji et al., 2013).

- **Ingresos y poder adquisitivo:** las variables “sueldos y salarios” y “población en edad de trabajar” presentan correlaciones positivas con el *Case-Shiller Index*, lo que indica que mayores ingresos y una población económicamente activa más amplia impulsan la demanda de vivienda y elevan sus precios.
- **Tasa de desempleo:** se observa una correlación negativa con el *Case-Shiller Index*, lo que sugiere que un menor desempleo aumenta la confianza del consumidor y estimula la demanda de vivienda.
- **Deuda hipotecaria:** muestra una fuerte relación con el *Case-Shiller Index*, lo que sugiere que una expansión del crédito hipotecario facilita la compra de viviendas y contribuye al alza en los precios del mercado inmobiliario.
- **Nuevas viviendas privadas en construcción e índice de producción industrial:** presentan correlaciones positivas con el *Case-Shiller Index*, indicando que la construcción y la actividad industrial están relacionadas con un mayor dinamismo del sector inmobiliario.
- **Precio del alquiler:** presenta una alta correlación positiva con el *Case-Shiller Index*, lo que sugiere que los aumentos en los precios de vivienda también elevan el coste del alquiler, reflejando una presión general sobre el mercado inmobiliario.

#### 4. MODELO DE CAMBIO DE RÉGIMEN DE MARKOV

##### 4.1. Introducción

El Modelo de Cambio de Régimen de Markov (*Markov-Switching Model*), propuesto por Hamilton (1989), es una herramienta econométrica ampliamente utilizada para detectar cambios estructurales en series temporales. También conocido como *regime-switching model*, es uno de los modelos no lineales más empleados en la literatura relativa al análisis de datos económicos y financieros.

Su principal ventaja radica en su capacidad para identificar distintos regímenes o estados en los datos y modelar las transiciones entre ellos mediante probabilidades condicionales. Al permitir cambios entre estas estructuras, el modelo logra capturar patrones dinámicos complejos, mejorando la capacidad de análisis en entornos con variaciones estructurales significativas (Chung-Ming, 2002).

En el ámbito inmobiliario, como se adelantó en el apartado anterior, diversos autores han empleado este modelo para la detección de burbujas especulativas, ya que permite segmentar el comportamiento de los precios de la vivienda en diferentes regímenes.

#### 4.2. Diseño e implementación

Siguiendo esta línea de investigación, en este trabajo se ha implementado un Modelo de Cambio de Régimen de Markov con el objetivo de identificar y analizar posibles burbujas inmobiliarias en el mercado de EE.UU. Nuestro enfoque se diferencia de estudios previos al centrarse en la aceleración del *Case-Shiller Index* en lugar de en su crecimiento o nivel absoluto. Esta decisión metodológica responde a la hipótesis de que los cambios en la dinámica del mercado se reflejan más claramente en la segunda derivada del precio que en su simple variación.

Para la implementación del modelo se ha utilizado un enfoque de tres regímenes, similar al propuesto por Nneji, Brooks y Ward (2013), con los siguientes estados:

Crecimiento estable ( <i>steady-state</i> )	Expansión acelerada ( <i>boom</i> )	Colapso de la burbuja inmobiliaria ( <i>crash</i> )
Fase en la que los precios de la vivienda evolucionan de manera sostenida sin señales de aceleración excesiva ni caídas abruptas	Fase caracterizada por un incremento acelerado en los precios, asociado a posibles burbujas especulativas	Período en el que los precios de la vivienda experimentan caídas pronunciadas, indicando un ajuste del mercado

*Tabla 2: Descripción estados modelo de cambio de régimen de Markov*

*Fuente: Elaboración propia*

Para cada uno de estos estados, se ha calculado la probabilidad de pertenencia en cada período. Este método permite una evaluación dinámica de la evolución del mercado inmobiliario, facilitando la identificación de transiciones entre regímenes.

En otro orden de cosas, se ha seleccionado un valor mínimo de probabilidad a partir del cual se considera que existe un riesgo significativo de estar en el régimen de colapso de burbuja inmobiliaria (*crash threshold*). Dicho valor ha sido fijado en 0.75, lo que implica que, cuando la probabilidad de pertenecer a este régimen supera este umbral, se activa una alerta de posible sobrevaloración en el mercado. La selección de este umbral

está sujeta a criterios de tolerancia al riesgo, que pueden ajustarse en función de los objetivos específicos del análisis.

## 5. MODELOS DE MACHINE LEARNING

A continuación, se han desarrollado algoritmos de *Machine Learning* con el objetivo de crear una herramienta predictiva que funcione como un mecanismo de alerta temprana. Para ello, se ha utilizado la probabilidad estimada por el Modelo de Cambio de Régimen de Markov, específicamente la probabilidad de estar en el régimen de colapso de la burbuja inmobiliaria (*crash*), como variable objetivo en modelos supervisados de predicción. Los modelos han sido entrenados con datos históricos de las variables explicativas recogidas en la Tabla 1, permitiendo identificar patrones y mejorar la capacidad de anticipación ante posibles episodios de sobrevaloración en el mercado inmobiliario.

### 5.1. Algoritmos empleados

#### 5.1.1. *Random Forest*

*Random Forest* es un algoritmo de aprendizaje supervisado basado en la construcción de múltiples árboles de decisión para mejorar la precisión y robustez de las predicciones.

Dentro de los árboles de decisión, durante su entrenamiento, las observaciones se dividen en subconjuntos a través de bifurcaciones en los nodos del árbol, hasta alcanzar un nodo terminal. De esta manera, cuando se quiere predecir una nueva observación, se recorre el árbol según el valor de sus predictores hasta llegar a un nodo terminal, donde la predicción corresponde a la media de la variable respuesta de las observaciones de entrenamiento en ese nodo. En el caso del *Random Forest*, cada árbol se entrena utilizando una muestra aleatoria con reemplazamiento del conjunto de datos original, técnica conocida como *bootstrapping*. Para realizar una predicción, se combinan las salidas de todos los árboles de modo que la predicción es el promedio de las predicciones de todos ellos (Amat, 2017).

Así, en la implementación de este algoritmo, el número de árboles de decisión es un hiperparámetro clave que debe seleccionarse adecuadamente. Un número bajo de árboles puede generar un modelo con mayor varianza y menor estabilidad, mientras que un

número excesivamente alto puede incrementar el coste computacional sin mejoras significativas en el rendimiento del modelo.

Aunque *Random Forest* mejora la capacidad predictiva en comparación con un único árbol de decisión, también presenta ciertas limitaciones. En particular, al combinar múltiples árboles, se reduce la interpretabilidad del modelo, lo que dificulta su representación visual y la identificación inmediata de los predictores más relevantes (Amat, 2017).

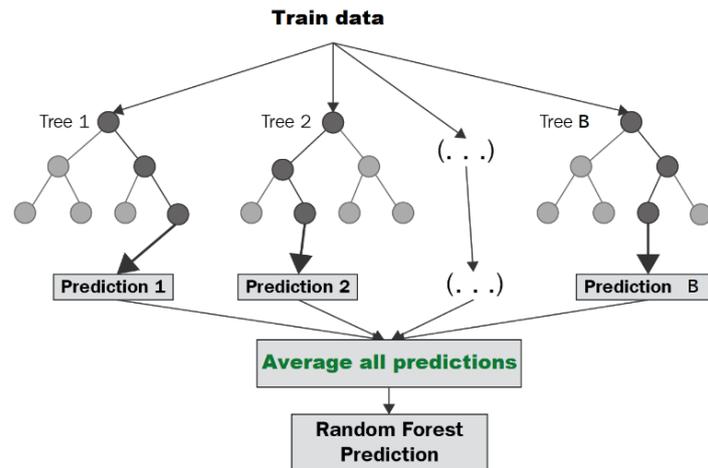


Imagen 7: Esquema del funcionamiento del algoritmo *Random Forest*

Fuente: Hernández, F. (2024)

### 5.1.2. XGBoost (*Extreme Gradient Boosting*)

Por su parte, XGBoost (*Extreme Gradient Boosting*) es un algoritmo que optimiza árboles de decisión de forma iterativa, mejorando la precisión del modelo al corregir los errores de predicciones previas (Chandra, 2023). Se basa en la técnica de *Gradient Boosting*, consistente en la construcción secuencial de árboles, donde cada nuevo árbol se entrena para minimizar los residuos del modelo anterior, ajustando los pesos de las observaciones mal predichas (Amat, 2017).

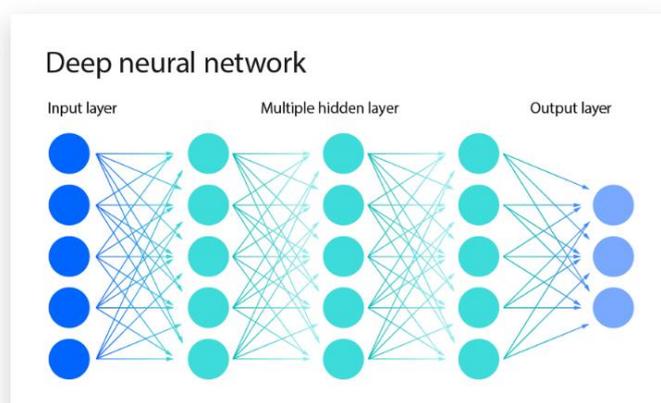
De igual manera que se mencionó previamente en relación con el algoritmo *Random Forest*, también para XGBoost el número de árboles de decisión es un hiperparámetro que ha de ser elegido cuidadosamente.

### 5.1.3. Redes neuronales artificiales

Las redes neuronales artificiales (*Artificial Neural Networks*, ANN) son los algoritmos de *Machine Learning* que constituyen la base de los modelos de aprendizaje profundo (*Deep Learning*). Su funcionamiento se inspira en el cerebro humano, ya que toman decisiones mediante procesos que imitan la forma en que las neuronas biológicas trabajan en conjunto para procesar información y reconocer patrones (IBM, 2024).

Cada red neuronal está compuesta por múltiples capas de nodos o neuronas artificiales, organizadas en tres tipos principales:

- **Capa de entrada (*input layer*):** recibe los datos iniciales del modelo.
- **Capas ocultas (*hidden layers*):** procesan la información aplicando transformaciones a través de conexiones ponderadas. Pueden ser una o varias, dependiendo de la complejidad del modelo.
- **Capa de salida (*output layer*):** genera la predicción o resultado final del modelo (IBM, 2024).



*Imagen 8: Esquema de funcionamiento del algoritmo ANN*

*Fuente: IBM (2024)*

Cada nodo dentro de la red está conectado con otros nodos y tiene asociados pesos y umbrales. Si la salida de un nodo supera un determinado valor umbral, ese nodo se activa y transmite la información a la siguiente capa de la red. En caso contrario, la señal no se propaga (IBM, 2024).

## 5.2. Diseño e implementación

Como paso previo común a la construcción de los tres algoritmos de *Machine Learning* predictivos y, a diferencia de los modelos explicativos, es necesario realizar una partición del conjunto de datos, proceso comúnmente conocido como *train-test split*.

El conjunto de entrenamiento (*training set*) se utilizará para ajustar el modelo, mientras que el conjunto de prueba (*test set*) permitirá comparar los valores predichos con los reales y cuantificar el error del modelo. En este estudio, la partición se ha establecido en un 80% para entrenamiento y 20% para validación.

Asimismo, para la selección óptima de los hiperparámetros de los dos primeros algoritmos, se ha llevado a cabo un proceso conocido como validación cruzada (*cross-validation*), el cual permite evaluar el rendimiento del modelo con distintos valores de hiperparámetros y seleccionar aquellos que maximizan la métrica de desempeño, en este caso, el coeficiente de determinación o  $R^2$ , cuyo significado se detallará más adelante en el trabajo (Omarzai, 2024).

De acuerdo con los resultados obtenidos a través de este procedimiento, el número óptimo de árboles de decisión (*n\_estimators*) para *Random Forest* es 500, mientras que para *XGBoost* es 100, valores que ofrecen el mejor equilibrio entre precisión y eficiencia computacional.

## V. ANÁLISIS DE RESULTADOS EMPÍRICOS

### 1. MODELO DE CAMBIO DE RÉGIMEN DE MARKOV

El Modelo de Cambio de Régimen de Markov aplicado en este estudio permite analizar la evolución del mercado inmobiliario estadounidense entre 1987 y 2024, segmentándolo en tres regímenes distintos: crecimiento estable (*steady-state*), expansión acelerada (*boom*) y colapso de la burbuja inmobiliaria (*crash*). A través de la estimación de las probabilidades de pertenencia de cada período a cada régimen, es posible identificar con precisión los momentos en los que el mercado inmobiliario ha experimentado fases de sobrevaloración y corrección de precios.

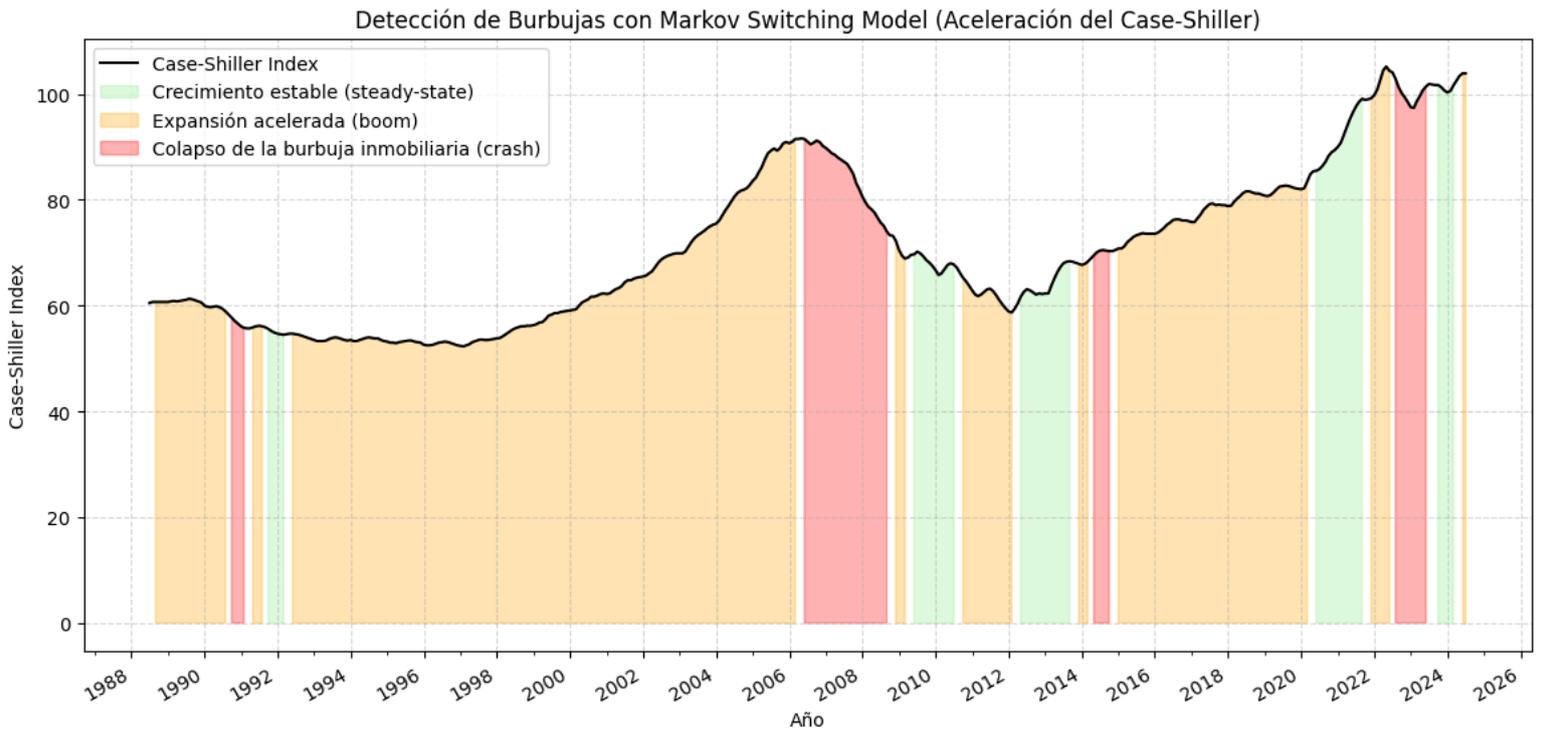


Imagen 9: Identificación de ciclos inmobiliarios según Markov Switching Model

Fuente: Elaboración propia en Python

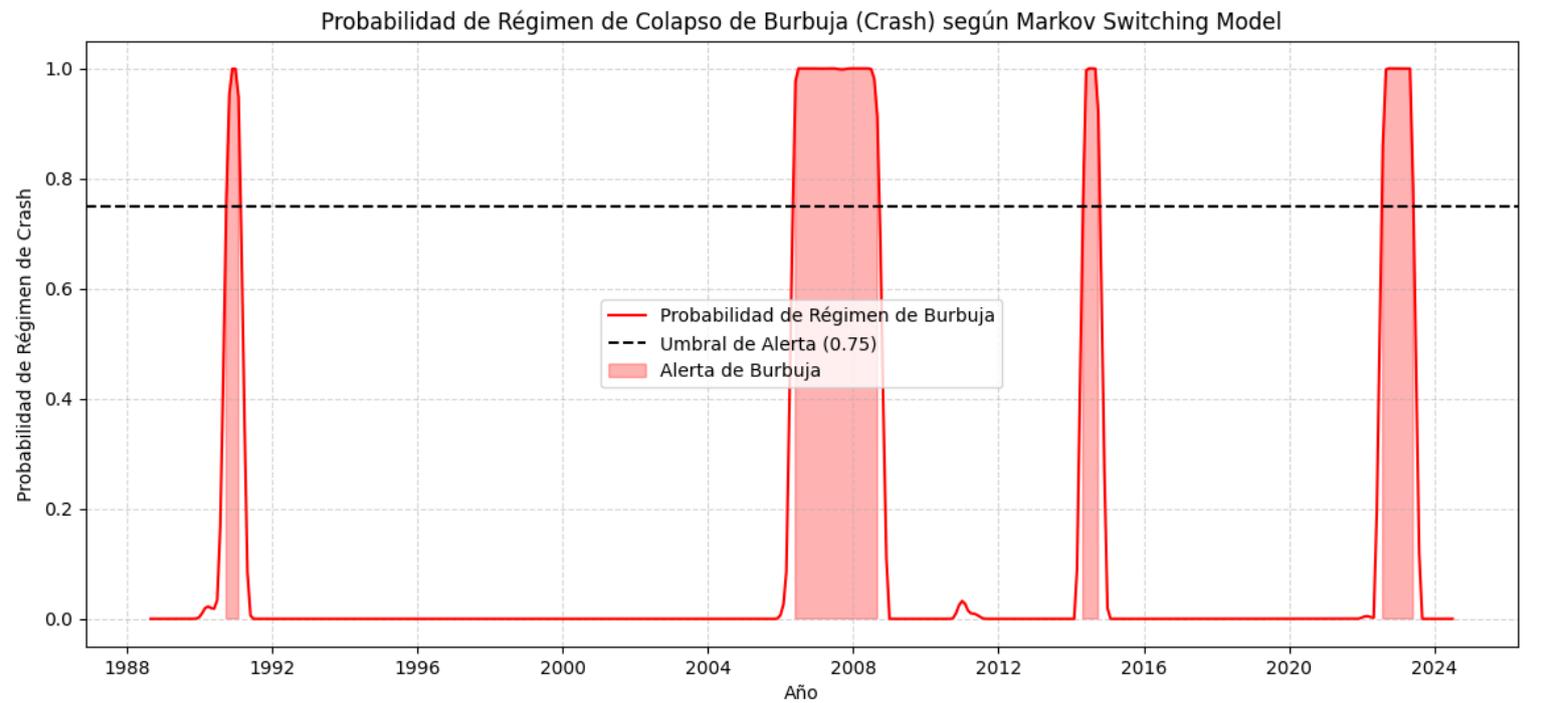


Imagen 10: Probabilidad de régimen de colapso de burbuja (crash) según Markov Switching Model

Fuente: Elaboración propia en Python

Los resultados empíricos obtenidos indican que el mercado inmobiliario ha atravesado varias fases críticas en las últimas décadas, lo que concuerda con la naturaleza cíclica del mercado inmobiliario identificada en la literatura. En la Imagen 9, se observa la evolución del *Case-Shiller Index*, junto con la clasificación de los períodos en los tres regímenes establecidos: crecimiento estable, expansión acelerada (*boom*) y colapso de la burbuja inmobiliaria (*crash*).

Por su parte, la Imagen 10 muestra la evolución de la probabilidad de que el mercado inmobiliario se encuentre en un régimen de colapso de la burbuja, utilizando un umbral (*crash threshold*) de 0.75 como criterio para definir la presencia de sobrevaloración.

### **1.1. Principios de los 90**

A finales de la década de 1980, el mercado inmobiliario estadounidense atravesó un episodio caracterizado por una rápida aceleración de los precios, seguida de una fuerte corrección a comienzos de los 90, resultado de una combinación de factores económicos y financieros adversos. En particular, la recesión de principios de los 90 fue impulsada por tres elementos clave.

En primer lugar, la política monetaria restrictiva implementada por la Reserva Federal, cuyo objetivo era controlar la inflación, resultó en un aumento de las tasas de interés y una reducción del crédito disponible para la compra de viviendas (Hoang, 2024).

Por otro lado, la Crisis de Ahorro y Préstamos (*Savings and Loan Crisis*) afectó a cientos de instituciones financieras en EE.UU., lo que redujo significativamente la disponibilidad de crédito hipotecario y generó incertidumbre en el mercado inmobiliario (Robinson, 2013).

Finalmente, el colapso del mercado de valores en 1987 provocó una crisis de confianza en los mercados financieros, afectando a la inversión y el consumo, lo que contribuyó al debilitamiento del sector inmobiliario (Bernhardt y Eckbald, 2013).

Como resultado de estos factores, los precios de la vivienda se estancaron durante los primeros años de la década, con una reducción en la demanda y un mercado caracterizado por baja liquidez y menor actividad en la construcción. A partir de mediados de los años 90, el mercado comenzó a recuperarse lentamente, impulsado por la reducción de las tasas de interés, en cuanto la Reserva Federal flexibilizó su política monetaria, facilitando el acceso al crédito. Además, la estabilización económica y la mejora en la

confianza de los consumidores reactivaron gradualmente la inversión inmobiliaria (Hoang, 2024).

Los resultados obtenidos en el Modelo de Cambio de Régimen de Markov reflejan esta evolución, mostrando un período de alta probabilidad de corrección del mercado en los primeros años de los 90, seguido de una fase de crecimiento estable a medida que el sector inmobiliario se recuperaba. Este hallazgo es consistente con los estudios de otros autores, como Fabozzi et al., quienes también identifican períodos de sobrevaloración en el mercado inmobiliario estadounidense a finales de la década de los 80 y principios de los 90 (Fabozzi et al., 2020).

## **1.2. Crisis inmobiliaria de las hipotecas *suprime***

A mediados de la década de los 2000, el mercado inmobiliario estadounidense experimentó una fase de crecimiento acelerado, impulsado por la expansión del crédito hipotecario, la titulización de activos financieros y la relajación de los estándares crediticios. Durante este período, el acceso a hipotecas se amplió de manera significativa, facilitando la compra de viviendas a prestatarios con historiales crediticios deficientes, lo que resultó en una sobrevaloración del mercado inmobiliario (Duca, 2013). Esta fase de crecimiento sostenido llevó a una burbuja inmobiliaria, que culminó en la crisis de las hipotecas *subprime* de 2007-2008, una de las crisis financieras más severas en la historia reciente de EE.UU.

El auge del mercado inmobiliario en los años 2000 se sustentó principalmente en la proliferación de préstamos hipotecarios *subprime*, dirigidos a prestatarios de alto riesgo. Si bien ello facilitó el acceso a la vivienda, muchas de estas hipotecas se encontraban sujetas a tipos de interés variables, lo que conllevaba la posibilidad de que los pagos mensuales se incrementaran drásticamente con el tiempo, afectando la capacidad de pago de los prestatarios y aumentando el riesgo de incumplimiento (Duca, 2013).

Tanto prestamistas como inversores asumieron que los precios de la vivienda continuarían aumentando indefinidamente, lo que llevó a una excesiva toma de riesgos y alimentó la especulación en el mercado. Al mismo tiempo, la falta de supervisión efectiva permitió la expansión de productos hipotecarios de alto riesgo, muchos de los cuales contenían cláusulas de pago insostenibles.

En resumen, aunque la expansión del crédito *subprime* facilitó la compra de vivienda para más de 5 millones de personas, sus beneficios quedaron eclipsados por los efectos negativos asociados a prácticas de préstamo abusivas, la rigidez en la modificación de los términos hipotecarios y el aumento de los incumplimientos por parte de los prestatarios (Jaffee, 2008).

Todo ello derivó en una sobrevaloración generalizada del mercado inmobiliario, que alcanzó su punto máximo en 2006 antes de experimentar un colapso pronunciado. Esto es consistente con los resultados empíricos obtenidos en el Modelo de Cambio de Régimen de Markov construido en este estudio. Como se observa en la Imagen 6, el modelo identifica un régimen de expansión acelerada (*boom*) a principios de los 2000, reflejando el crecimiento exponencial de los precios impulsado por el crédito y la especulación. Posteriormente, entre 2006 y 2008, se detecta una transición hacia el régimen de colapso de la burbuja inmobiliaria (*crash*), coincidiendo con el estallido de la crisis *subprime* y la fuerte corrección de precios que afectó al mercado.

Estos resultados son consistentes con los hallazgos de otros estudios, en particular los de Carlos Cañizares, quien, a través de su modelo de Markov, también identifica una alta probabilidad de estallido de burbuja durante la Gran Recesión a finales de la década de 2000 (Cañizares Martínez, 2023).

### **1.3. Post-crisis y recuperación gradual**

Tras el colapso inmobiliario de 2008 y la crisis financiera global, el mercado de la vivienda en EE.UU. experimentó una recuperación progresiva (Hermann y Whitney, 2024). El modelo de Cambio de Régimen de Markov construido muestra que, durante este período, el mercado pasó por una fase de crecimiento estable (*steady-state*) con algunas fluctuaciones, alternando entre periodos de estabilidad y expansión. Como factores clave que pudieron contribuir a esta tendencia, la Reserva Federal mantuvo tasas de interés cercanas a cero, facilitando el acceso al crédito hipotecario y estimulando la demanda de viviendas. Asimismo, la implementación de medidas más estrictas para la concesión de hipotecas evitó la formación de otra burbuja inmobiliaria a corto plazo. Con todo ello, la recuperación del mercado laboral y el incremento de los ingresos familiares ayudaron a reforzar la estabilidad del sector inmobiliario.

No obstante, la Imagen 6 también sugiere picos de aceleración (*boom*) en ciertos momentos. En este sentido, para 2015, los precios de las viviendas habían aumentado significativamente, acercándose a los niveles previos a la crisis en algunas áreas metropolitanas. Este incremento generó debates sobre la posibilidad de una nueva burbuja inmobiliaria. Sin embargo, la mayoría de los economistas coincidía en que, aunque los precios estaban en alza, el mercado presentaba fundamentos más sólidos en comparación con la década anterior (Vasel, 2015).

#### **1.4. Impacto de la pandemia de COVID-19**

La pandemia de COVID-19 desencadenó un aumento sin precedentes en los precios de la vivienda en EE.UU., marcando un período de crecimiento acelerado que superó cualquier otro en la historia reciente del mercado inmobiliario. Durante estos años, se establecieron y superaron récords en el crecimiento de los precios, reflejando una dinámica de mercado excepcional.

Este incremento se debió a una combinación de factores. Las tasas de interés históricamente bajas establecidas por la Reserva Federal facilitaron el acceso al crédito y alentaron la compra de viviendas. Simultáneamente, la pandemia transformó las preferencias de los compradores, aumentando la demanda de propiedades con mayores espacios interiores y exteriores debido al confinamiento y la adopción masiva del teletrabajo. Otro factor clave fue el exceso de liquidez en el mercado. Durante la pandemia, muchas familias y compradores potenciales acumularon ahorros, ya que los gastos en servicios no esenciales se redujeron drásticamente. Esto les permitió permitirse el pago de precios más altos por las viviendas disponibles, intensificando la competencia en el mercado y elevando los precios aún más (Hermann y Whitney, 2024).

Al mismo tiempo, este incremento de la demanda se encontró con una oferta de vivienda limitada. Tras más de una década de construcción insuficiente, la escasez estructural de viviendas se hizo aún más evidente, lo que generó una presión significativa sobre los precios y un entorno de sobrevaloración en muchas regiones del país (Hermann y Whitney, 2024).

Los resultados obtenidos en el Modelo de Cambio de Régimen de Markov reflejan esta dinámica, mostrando una fase prolongada de expansión en el mercado inmobiliario entre 2020 y 2022. Sin embargo, a partir de 2023, el aumento progresivo de las tasas de

interés por parte de la Reserva Federal y la creciente preocupación por la sostenibilidad de los precios comenzaron a desacelerar el crecimiento del mercado, lo que podría indicar el inicio de un ajuste en los valores de la vivienda (Hermann y Whitney, 2024).

### 1.5. Tendencias actuales

A pesar del acelerado aumento en los precios de la vivienda en los últimos años, los expertos coinciden en que el mercado inmobiliario se encuentra en un proceso de corrección gradual en lugar de un colapso abrupto, lo que hace improbable una caída de los precios similar a la vivida en la crisis de 2008 (Rothstein, 2025).

Según los especialistas, el repunte de los precios en 2021 no respondió a la especulación ni al crédito desmedido, sino a tasas hipotecarias históricamente bajas y a un cambio estructural en la demanda de vivienda, derivado de la pandemia y las nuevas preferencias de los compradores (Rothstein, 2025).

Otro aspecto que diferencia la situación actual de la crisis de 2008 es la mayor estabilidad del sistema financiero. Actualmente, las normas para la concesión de préstamos son mucho más estrictas, gracias a las regulaciones impuestas tras la crisis *subprime*, lo que ha reducido considerablemente el riesgo de impago masivo. Además, la mayoría de los propietarios actuales cuenta con hipotecas a tasa fija, lo que significa que, aunque los tipos de interés hayan aumentado significativamente, sus pagos mensuales se han mantenido estables (Rothstein, 2025).

En definitiva, aunque el auge inmobiliario impulsado por la pandemia fue intenso, las políticas macroprudenciales contribuyeron a contener el crédito, prevenir un sobrecalentamiento del mercado y mitigar los riesgos financieros, lo que permitió desinflar la burbuja a tiempo y mantener la estabilidad del sistema bancario y financiero (Martínez García, 2024). En este contexto, el riesgo actual de un colapso inmobiliario en EE.UU. es relativamente bajo, siendo más probable que el mercado atraviese una desaceleración progresiva o una corrección moderada de precios en los próximos años.

## 2. EVALUACIÓN Y COMPARACIÓN DE LOS ALGORITMOS DE *MACHINE LEARNING*

Una vez examinados los resultados del Modelo de Cambio de Régimen de Markov, resulta fundamental evaluar el desempeño de los algoritmos de *Machine Learning*

desarrollados. Para ello, se comparará su rendimiento en función de diversas métricas, con el objetivo de identificar el modelo óptimo como herramienta de alerta temprana.

## 2.1. Descripción de métricas de evaluación empleadas

En el ámbito de la evaluación de los modelos de *Machine Learning* se emplean diversos indicadores para cuantificar la precisión y eficacia de las predicciones. A continuación, se describen las principales métricas utilizadas.

### 2.1.1. Error absoluto medio (*Mean Absolute Error, MAE*)

En primer lugar, el error absoluto medio (MAE) mide la magnitud promedio de los errores en un conjunto de predicciones, sin considerar su dirección. Se calcula como la media de las diferencias absolutas entre los valores predichos y los valores reales (Kumar, 2024).

En este caso, el MAE equivale a la diferencia entre la probabilidad de colapso de la burbuja inmobiliaria (*crash*) identificada por el Modelo de Cambio de Régimen de Markov y la probabilidad predicha por el algoritmo de *Machine Learning*.

### 2.1.2. Error cuadrático medio (*Mean Squared Error, MSE*) y raíz del error cuadrático medio (*Root Mean Squared Error, RMSE*)

En segundo lugar, el error cuadrático medio (MSE) es una métrica que calcula el promedio de los errores al cuadrado, lo que permite evaluar la precisión del modelo enfatizando los errores más grandes.

Al elevarse los errores al cuadrado, los valores atípicos tienen un impacto considerablemente mayor en el resultado final, lo que convierte al MSE en una métrica más sensible a los errores extremos en comparación con el MAE (Kumar, 2024).

Por su parte, el RMSE corresponde a la raíz cuadrada del MSE, lo que introduce la posibilidad de interpretar el error en las mismas unidades que la variable objetivo.

### 2.1.3. Error cuadrático medio ponderado (*Weighted Mean Squared Error, WMSE*)

El error cuadrático medio ponderado (WMSE) es una variante del MSE que asigna diferentes pesos a cada observación, permitiendo enfatizar ciertas áreas del conjunto de datos.

En este estudio, se ha considerado esencial utilizar el WMSE con mayores pesos asignados a períodos con alta probabilidad de colapso de la burbuja inmobiliaria. De este modo, los errores cometidos en estas fases tienen un impacto mayor en la evaluación del modelo, permitiendo un análisis más detallado del rendimiento en los momentos más relevantes del ciclo económico.

### 2.1.4. Coeficiente de determinación ( $R^2$ )

El coeficiente de determinación ( $R^2$ ) mide la proporción de la variabilidad de la variable objetivo (en este caso, la probabilidad de estar en un régimen de colapso de la burbuja inmobiliaria) que es explicada por el modelo. Su valor varía entre 0 y 1, donde valores más cercanos a 1 indican un mejor ajuste. De esta manera, un  $R^2$  alto indica que el modelo puede explicar gran parte de la variabilidad observada en los datos, mientras que un valor bajo sugiere que el modelo no se ajusta bien a los datos (Kumar, 2024).

### 2.1.5. Correlación de Pearson

Finalmente, el coeficiente de correlación de Pearson mide la fuerza y dirección de la relación lineal entre dos variables. El coeficiente de Pearson varía entre -1 y 1, siendo 1 una correlación positiva perfecta, -1 una correlación negativa perfecta y 0 la ausencia de correlación lineal.

## 2.2. Análisis comparativo del desempeño de los algoritmos

Tras evaluar los modelos de *Random Forest*, XGBoost y Redes Neuronales Artificiales (ANN) mediante las métricas de desempeño previamente descritas, los resultados obtenidos permiten identificar cuál de ellos presenta mayor precisión y capacidad de generalización en la predicción de la probabilidad de colapso de la burbuja inmobiliaria.

	MAE	MSE	RMSE	WMSE	R <sup>2</sup>	Correlación de Pearson
Random Forest	0.04	0.01	0.10	0.04	0.80	0.90
XGBoost	0.04	0.01	0.12	0.06	0.71	0.84
ANN	0.01	~0.00	0.02	~0.00	0.97	0.99

Tabla 3: Resumen métricas de evaluación de los modelos de Machine Learning

Fuente: Elaboración propia

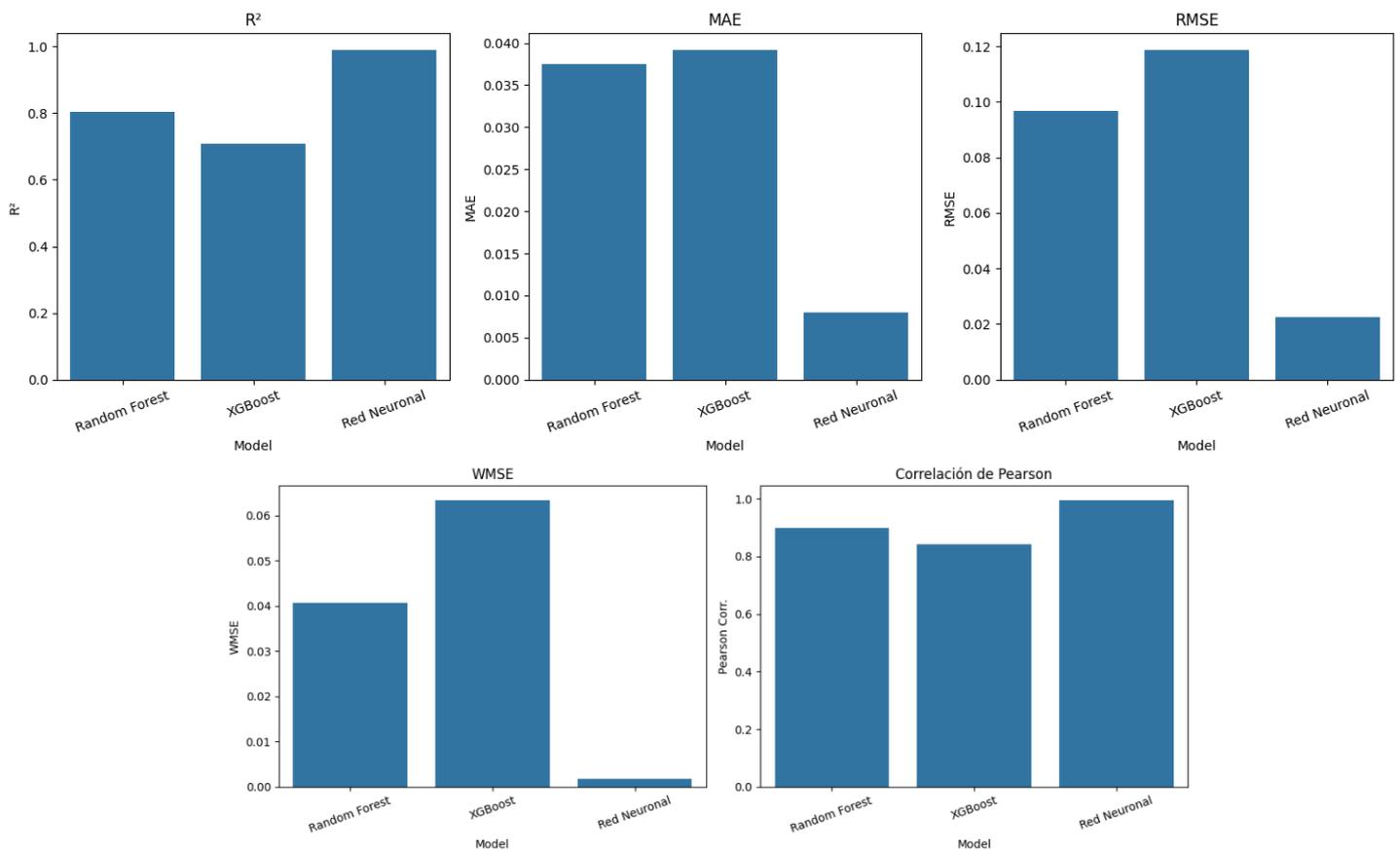
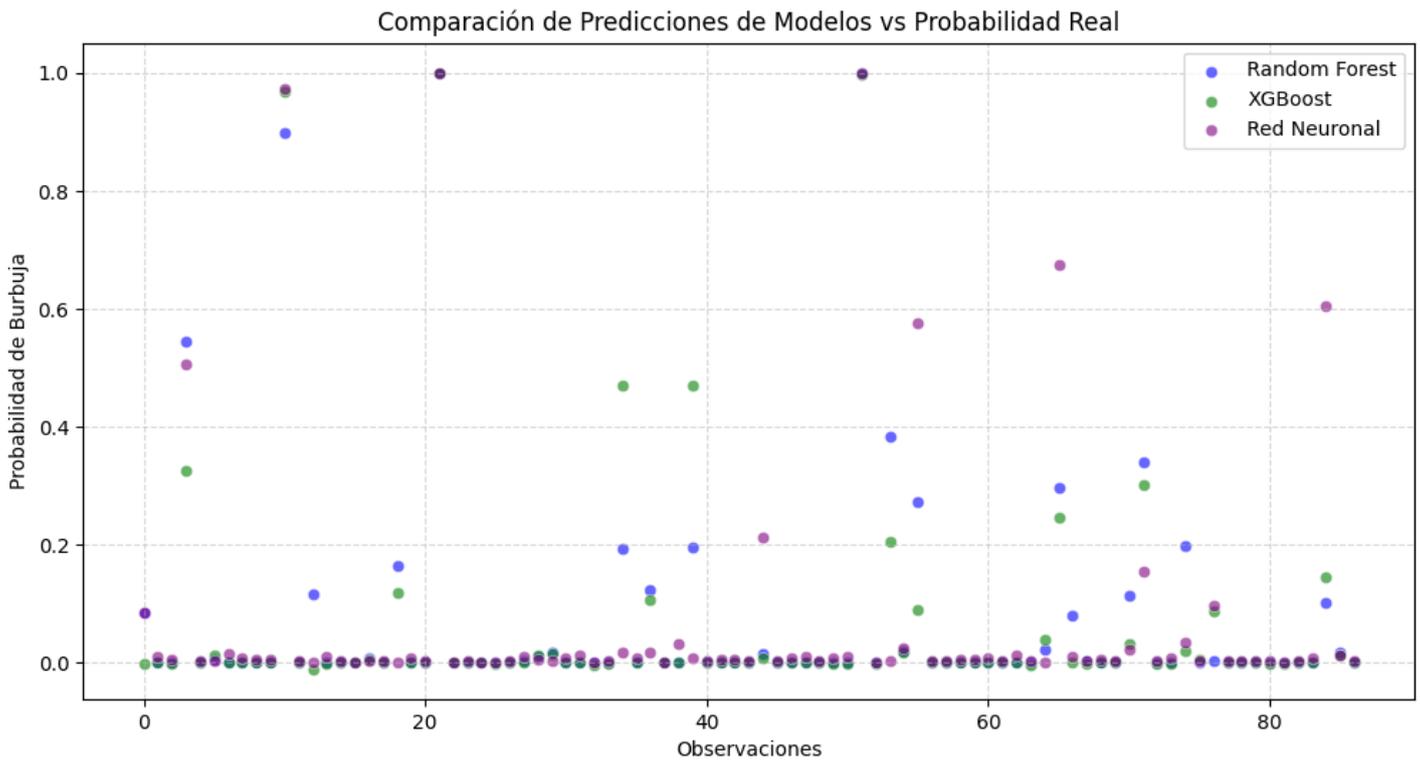


Imagen 11: Comparación de métricas de evaluación de los modelos de Machine Learning

Fuente: Elaboración propia en Python



*Imagen 12: Scatter plot comparativo de las predicciones de los modelos de Machine*

*Fuente: Elaboración propia en Python*

Como puede observarse en la Tabla 3 y la Imagen 11, el modelo de Red Neuronal Artificial (ANN) se posiciona como el modelo con mejor desempeño en todas las métricas. Presenta el menor MAE (0.01) y MSE ( $\sim 0.00$ ), lo que indica que sus errores absolutos y cuadráticos medios son significativamente más bajos en comparación con los otros modelos. Asimismo, su RMSE (0.02) es el más reducido, reflejando una menor dispersión en las predicciones respecto a los valores reales.

En términos del coeficiente de determinación ( $R^2$ ), la ANN alcanza un valor de 0.97, lo que sugiere que explica el 97% de la variabilidad en la probabilidad de colapso de la burbuja inmobiliaria, superando ampliamente a *Random Forest* (0.80) y XGBoost (0.71). Además, presenta la mayor correlación de Pearson (0.99), lo que indica una relación más fuerte entre sus predicciones y los valores reales.

Por otro lado, el modelo de *Random Forest* muestra un desempeño competitivo, con valores de MAE (0.04), MSE (0.01) y RMSE (0.10) inferiores a los de XGBoost, lo que sugiere que ofrece una mayor precisión que este último. Su  $R^2$  de 0.80 y su correlación de Pearson de 0.90 indican que, aunque no alcanza la precisión de la ANN, también es un modelo sólido con buen desempeño.

Finalmente, XGBoost, aunque eficiente, presenta un rendimiento inferior en comparación con los otros dos modelos. Su MSE (0.01) y RMSE (0.12) reflejan una mayor dispersión en las predicciones, mientras que su  $R^2$  de 0.71 sugiere que explica menos variabilidad en la probabilidad de colapso de la burbuja que los otros modelos.

Un aspecto relevante a considerar es la diferencia entre el MSE y el WMSE, ya que en todos los modelos el WMSE es mayor que el MSE. Esto indica que los modelos tienen un desempeño relativamente menor en la predicción de períodos de alta probabilidad de colapso de la burbuja en comparación con aquellos de baja probabilidad. Esta diferencia es especialmente importante, ya que los períodos de alta probabilidad de *crash* son los más relevantes para la detección de crisis inmobiliarias y la formulación de alertas tempranas. Sin embargo, el WMSE sigue siendo relativamente bajo, lo que implica que, en términos generales, los modelos mantienen una capacidad predictiva adecuada incluso en los escenarios más críticos.

Con base en estos resultados, se concluye que la Red Neuronal Artificial (ANN) es el modelo más adecuado para la predicción de la probabilidad de colapso de la burbuja inmobiliaria, dada su mayor precisión y capacidad explicativa.

## VI. CONCLUSIONES

En este Trabajo de Fin de Grado se ha explorado la integración de Modelos de Cambio de Régimen de Markov y algoritmos de *Machine Learning* como herramientas para la detección de burbujas inmobiliarias.

Los resultados empíricos reflejan que el mercado inmobiliario sigue patrones cíclicos, donde períodos de crecimiento sostenido pueden derivar en fases de exuberancia, seguidas de ajustes y colapsos. El uso de técnicas avanzadas ha permitido identificar correctamente momentos históricos clave, como la burbuja *subprime* de 2008 y el rápido repunte de precios posterior a la pandemia de COVID-19. Asimismo, se ha demostrado que los modelos predictivos pueden ser herramientas valiosas para inversores, analistas y reguladores en la toma de decisiones estratégicas.

En consecuencia, se presentan a continuación posibles aplicaciones de los mecanismos de alerta construidos, así como limitaciones potenciales del estudio que pueden ser abordadas en futuras investigaciones para mejorar la precisión y aplicabilidad de los modelos desarrollados.

## 1. PROPUESTAS DE APLICACIÓN

Los modelos desarrollados en este estudio tienen un amplio potencial de aplicación en diversos ámbitos, desde la regulación financiera y la política de vivienda hasta la inversión inmobiliaria.

En el ámbito regulatorio, bancos centrales y entidades gubernamentales podrían utilizar estos mecanismos de alerta para adoptar medidas “macroprudenciales” y monitorear la estabilidad financiera, ajustando políticas de crédito y tasas de interés en función del riesgo de burbuja presente en cada momento. Además, su integración en la planificación urbana y política de vivienda permitiría detectar desequilibrios regionales y diseñar estrategias para garantizar la sostenibilidad del mercado inmobiliario.

Desde otro punto de vista, los inversores podrían apalancarse en este tipo de herramientas para identificar señales tempranas de sobrevaloración del mercado, optimizando estrategias de compra y venta, así como en la gestión de carteras inmobiliarias y fondos de inversión.

En definitiva, la integración de estos modelos en herramientas de análisis permitiría mejorar la toma de decisiones estratégicas, anticipar riesgos y contribuir a la estabilidad del mercado inmobiliario a nivel global.

## 2. POTENCIALES LIMITACIONES

A pesar de las múltiples ventajas que brinda la combinación del Modelo de Cambio de Régimen de Markov y los algoritmos de *Machine Learning* propuesta, es importante mencionar ciertas limitaciones que deben considerarse en la interpretación de los resultados empíricos obtenidos.

En primer lugar, si bien este trabajo propone una posible metodología aplicable a la detección de burbujas inmobiliarias, es importante mencionar ciertas limitaciones relevantes relacionadas con la predicción tardía de la existencia de una burbuja. En particular, los modelos construidos tienden a estimar la probabilidad de caídas pronunciadas en los precios cuando el ajuste del mercado (*crash*) ya está en marcha. Esta demora reduce en parte la utilidad práctica del modelo para decisiones preventivas, lo que evidencia la necesidad de seguir mejorando los mecanismos de detección temprana, sofisticando los indicadores para anticipar con la mayor antelación posible el cambio de régimen del mercado.

Por otro lado, la selección de tres regímenes es una decisión metodológica basada en estudios previos, pero no garantiza que el *regime-switching* model de Markov refleje con exactitud todos los estados posibles del mercado. En este sentido, es posible que existan transiciones más graduales o regímenes adicionales que no han sido capturados en este enfoque.

Asimismo, la elección del umbral de colapso (*crash threshold*) de 0.75 es en cierto modo arbitraria y su impacto en la detección de burbujas podría variar según la tolerancia al riesgo del analista en cada caso. Una sensibilidad distinta en este umbral podría generar diferencias en la identificación de períodos de sobrevaloración del mercado.

Otro aspecto a considerar es el hecho de que el Modelo de Cambio de Régimen de Markov parece predecir una alerta de burbuja inmobiliaria en períodos en los que, pese a existir un crecimiento acelerado del precio de la vivienda, dicho crecimiento podría estar justificado por fundamentos económicos sólidos y no necesariamente por factores especulativos. Esto sugiere la necesidad de complementar el modelo con indicadores macroeconómicos adicionales, como el nivel de endeudamiento de los hogares, con el fin de mejorar la precisión en la detección de burbujas. En esta línea, Carlos Cañizares (2023) propone en su estudio la incorporación de variables relacionadas con el nivel de endeudamiento hipotecario para mejorar significativamente la capacidad predictiva del modelo.

En lo que respecta a los algoritmos de *Machine Learning*, es importante recordar que aprenden patrones a partir de datos históricos, pero no siempre son capaces de anticipar eventos inesperados o cambios estructurales en el mercado. Además, existe el denominado riesgo de sobreajuste (*overfitting*), en el que los modelos pueden ajustarse excesivamente a los datos de entrenamiento, reduciendo su capacidad de generalización y afectando a su desempeño en datos nuevos.

Asimismo, una limitación adicional a considerar radica en la naturaleza acotada de la variable dependiente utilizada en este estudio, la cual toma valores continuos en el intervalo  $[0,1]$ . En este sentido, la literatura advierte que muchos modelos tradicionales de *Machine Learning* asumen que la variable objetivo puede tomar cualquier valor en la recta real y que los errores están normalmente distribuidos con varianza constante. Como señalan Smithson y Verkuilen (2006), aplicar estos modelos bajo tales condiciones puede conducir a interpretaciones erróneas. Para mitigar este problema, los autores proponen el

uso de modelos basados en la distribución beta, los cuales permiten modelar de forma explícita tanto la media como la dispersión, respetando al mismo tiempo los límites del intervalo unitario. Esta alternativa representa un potencial paso siguiente para perfeccionar la metodología propuesta en este trabajo.

Por último, el entrenamiento de modelos complejos de *Machine Learning*, en particular el de Redes Neuronales Artificiales (ANN), requiere una capacidad computacional y un tiempo de ajuste considerable, lo que puede representar un desafío en términos de eficiencia y disponibilidad de recursos, especialmente en entornos con restricciones de *hardware*.

## VII. DECLARACIÓN DEL USO DE HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL GENERATIVA

**ADVERTENCIA:** Desde la Universidad consideramos que ChatGPT u otras herramientas similares son herramientas muy útiles en la vida académica, aunque su uso queda siempre bajo la responsabilidad del alumno, puesto que las respuestas que proporciona pueden no ser veraces. En este sentido, NO está permitido su uso en la elaboración del Trabajo fin de Grado para generar código porque estas herramientas no son fiables en esa tarea. Aunque el código funcione, no hay garantías de que metodológicamente sea correcto, y es altamente probable que no lo sea.

Por la presente, yo, Claudia Gonzalo González, estudiante del Doble Grado en Análisis de Negocios/Business Analytics y Derecho (E-3 Analytics) de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado "Aplicación de *Machine Learning* para la detección de burbujas inmobiliarias", declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación:

1. **Brainstorming de ideas de investigación:** Utilizado para idear y esbozar posibles áreas de investigación.
2. **Metodólogo:** Para descubrir métodos aplicables a problemas específicos de investigación.
3. **Interpretador de código:** Para realizar análisis de datos preliminares.
4. **Corrector de estilo literario y de lenguaje:** Para mejorar la calidad lingüística y estilística del texto.
5. **Sintetizador y divulgador de libros complicados:** Para resumir y comprender literatura compleja.
6. **Revisor:** Para recibir sugerencias sobre cómo mejorar y perfeccionar el trabajo con diferentes niveles de exigencia.
7. **Traductor:** Para traducir textos de un lenguaje a otro.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado ChatGPT u otras herramientas similares). Soy

consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: 20/03/2025

Firma: Claudia Gonzalo González

## VIII. BIBLIOGRAFÍA

- Álvarez-Lois, P. & Nuño-Barrau, G. (2007). *The Role of Fundamentals in the Price of Housing: Theory and Evidence*. BBVA Economic Research Department, Working paper No. 0701. (disponible en [The Role of Fundamentals in the Price of Housing: Theory and Evidence | BBVA Research](#); última consulta 19/03/2025)
- Álvarez-Román, L. & García-Posada, L. (2021). *Are house prices overvalued in Spain? A regional approach*. *Economic Modelling*, Volume 99, 105499. (disponible en [Are house prices overvalued in Spain? A regional approach](#); última consulta 18/03/2025)
- Amat, J. (2017). *Árboles de decisión, random forest, grading boosting y C5.0*. *Ciencia de Datos*. (disponible en [Árboles de decisión, Random Forest, Gradient Boosting y C5.0](#); última consulta 19/03/2025)
- Ayuso, J. & Restoy, F. (2006). *El precio de la vivienda en España: ¿es robusta la evidencia de sobrevaloración?* Banco de España, Boletín Económico, Artículos Analíticos, pp. 59-65. (disponible en [El precio de la vivienda en España: ¿es robusta la evidencia de sobrevaloración?](#); última consulta 10/02/2025)
- Bernhardt, D. & Eckbald, M. (2013). *Stock Market Crash of 1987*. Federal Reserve History. (disponible en [Stock Market Crash of 1987 | Federal Reserve History](#); última consulta 15/03/2025)
- Bokreta, R. & Leblanc, M. (2009). *Analysis of the US Real Estate Market: Time-Varying Estimation and Forecast of the S&P Case-Shiller Composite 20 Cities*. SSRN Electric Journal. (disponible en [Analysis of the US Real Estate Market: Time-Varying Estimation and Forecast of the S&P Case-Shiller Composite 20 Cities](#); última consulta 19/03/2025)
- Cañizares Martínez, C. (2023). *Leaning against housing booms fueled by credit*. University of Milan Bicocca Department of Economics, Management and Statistics Working Paper No. 513. (disponible en [Leaning against housing booms fueled by credit](#); última consulta 19/03/2025)
- Chandra, R. (2023). *XGBoost: la potencia de los algoritmos de aprendizaje automático*. Kanaries. (disponible en [¿Qué es XGBoost, la potencia de los algoritmos de aprendizaje automático?](#); última consulta 15/03/2025)

- Coleman, M., LaCour-Little, M., & Vandell, K. (2008). *Subprime Lending and the Housing Bubble: Tail Wags Dog?* Journal of Housing Economics, Volume 17, Issue 4. (disponible en [Subprime Lending and the Housing Bubble: Tail Wags Dog?](#); última consulta 15/03/2025)
- Dreger, C. & Kholodilin, A. (2011). *An early warning system to predict the house price bubbles*. German Institute for Economic Research Discussion Papers, No. 114. (disponible en [An Early Warning System to Predict Speculative House Price Bubbles](#) última consulta 19/03/2025)
- Duca, J. (2013). *Subprime Mortgage Crisis*. Federal Reserve History. (disponible en [Subprime Mortgage Crisis | Federal Reserve History](#); última consulta 15/03/2025)
- Fabozzi, F., Kynigakis, I., Panopoulou, E. & Tunaru, R. (2020). *Detecting Bubbles in the US and UK Real Estate Markets*. SSRN Electric Journal. (disponible en [Detecting Bubbles in the US and UK Real Estate Markets](#); última consulta 19/03/2025)
- Federal Reserve Bank of St. Louis. (2024). *Federal Reserve Economic Data*. St. Louis FED. (disponible en [Federal Reserve Economic Data | FRED | St. Louis Fed](#); última consulta 1/03/2025)
- García-Montalvo, J. (2006). *Deconstruyendo la burbuja: expectativas de revalorización y precio de la vivienda en España*. La vivienda: precios, mercados y financiación. Papeles de Economía Española, No. 109, pp. 44-75. (disponible en [Are house prices overvalued in Spain? A regional approach](#); última consulta 15/03/2025)
- Hermann, A. & Whitney, P. (2024). *The Geography of Pandemic Era Home Price Trends and the Implications for Affordability*. Joint Center for Housing Studies of Harvard University. (disponible en [The Geography of Pandemic-Era Home Price Trends and the Implications for Affordability](#) última consulta 19/03/2025)
- Hernández, F. (2024). *Modelos Predictivos*. Github. (disponible en [Modelos Predictivos](#); última consulta 15/02/2025)
- Hoang, T. (2024). *Time series analysis on predicting the US housing market prices*. SSRN. (disponible en [Time series analysis on predicting the US housing market prices](#); última consulta 19/03/2025)

- IBM. (2024). *What is a neural network?* IBM Think. (disponible en [What is a Neural Network? | IBM](#); última consulta 17/03/2025)
- Jaffee, M. (2008). *The U.S. subprime mortgage crisis: issues raised and lessons learned*. World Bank. Commission on Growth and Development. Working Paper No. 28. (disponible en [The U.S. subprime mortgage crisis: issues raised and lessons learned](#); última consulta 15/03/2025)
- Jones, K. (2025). *Regime Switching Models for Time Series Analysis in Python*. Medium. (disponible en [Regime Switching Models for Time Series Analysis in Python](#); última consulta 15/02/2025)
- Chung-Ming, K. (2002). *Lecture on the Markov Switching Model*. Institute of Economics Academia Sinica. (disponible en [Lecture on the Markov Switching Model](#); última consulta 15/02/2025)
- Kumar, A. (2024). *MSE vs RMSE vs MAE vs MAPE vs R-Squared: When to Use?* Analytics Yogi. (disponible en [MSE vs RMSE vs MAE vs MAPE vs R-Squared: When to Use?](#); última consulta 15/03/2025)
- Martínez García, E. (2024). *How to spot housing bubbles*. Finance & Development Magazine, The Economics of Housing, International Monetary Fund, pp. 44-47. (disponible en [How To Spot Housing Bubbles](#); última consulta 14/02/2025)
- Mesquita, D. (2022). *Python AI: How to Build a Neural Network & Make Predictions*. Real Python. (disponible en [Python AI: How to Build a Neural Network & Make Predictions](#); última consulta 15/02/2025)
- Robinson, K. (2013). *Savings and Loan Crisis*. Federal Reserve History. (disponible en [Savings and Loan Crisis | Federal Reserve History](#); última consulta 15/03/2025)
- Rothstein, R. (2025). *Will The Housing Market Crash in 2024? Experts Give 5-Year Predictions*. Forbes. (disponible en [Will The Housing Market Crash in 2024?](#); última consulta 15/03/2025)
- Nneji, O., Brooks, C., & Ward, C. (2013). *House Price Dynamics and Their Reaction to Macroeconomic Changes*. Economic Modelling, Forthcoming. (disponible en [House Price Dynamics and Their Reaction to Macroeconomic Changes](#); última consulta 15/03/2025)

- Omarzai, F. (2024). *Hyperparameters Tuning And Cross Validation In Depth*. Medium. (disponible en [Hyperparameters Tuning And Cross Validation In Depth](#); última consulta 12/02/2025)
- Smithson, M., & Verkuilen, J. (2006). *A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables*. *Psychological Methods*, 11(1), 54–71. (disponible en <https://doi.org/10.1037/1082-989X.11.1.54>; última consulta 26/03/2025)
- S&P Dow Jones Indices. (2024). *S&P CoreLogic Case-Shiller Home Price Indices. Methodology*. S&P Global. (disponible en [S&P CoreLogic Case-Shiller Home Price Indices Methodology](#); última consulta 15/02/2025)
- Tuychiev, B. (2022). *Using XGBoost in Python Tutorial*. Datacamp. (disponible en [Learn XGBoost in Python: A Step-by-Step Tutorial](#); última consulta 15/02/2025)
- Ujhelyi, T. (2022). *Random Forest in Python (and coding it with Scikit-learn)*. Data36. (disponible en [Random Forest in Python and coding it with Scikit-learn](#); última consulta 15/02/2025)
- U.S. Bureau of Labor Statistics. (2021). *Unemployment rises in 2020, as the country battles the COVID-19 pandemic*. *Monthly Labor Review*. (disponible en [Unemployment rises in 2020, as the country battles the COVID-19 pandemic](#); última consulta 15/03/2025)
- Vasel, K. (2015). *Here's what's driving up housing prices*. CNN Money, Real Estate Special Report. (disponible en [Is it time to worry about another housing bubble?](#); última consulta 15/03/2025)
- Zhang, Q. (2021). *Housing price prediction based on multiple linear regression*. *Hindawi Scientific Programming*, vol. 2021, 9 pages. (disponible en [Housing Price Prediction Based on Multiple Linear Regression](#); última consulta 2/02/2025)

## **IX. ANEXOS**

### **1. BASE DE DATOS DE EXCEL**

Enlace a la base de datos de Excel construida (link a archivo OneDrive):

[database\\_TFG\\_vF.xlsx](#)

### **2. CÓDIGO DE PYTHON**

Enlace al código de Python desarrollado en la plataforma Google Colab:

[Código\\_Python\\_TFG\\_Burbuja\\_Inmobiliaria\\_EEUU\\_CGG\\_vF.ipynb](#)