



GRADO EN INGENIERÍA EN TECNOLOGÍAS INDUSTRIALES

TRABAJO FIN DE GRADO

MODELO PREDICTIVO DE INVERSIÓN EN OBRAS DE ARTE

Ortega Isasi, Eduardo

Director: Dr. Antonio García y de Garmendia

Madrid

Julio 2021

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título

Modelo Predictivo de Inversión en Obras de Arte

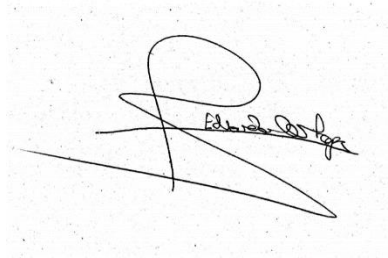
en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el

curso académico 2020/2021 es de mi autoría, original e inédito y

no ha sido presentado con anterioridad a otros efectos.

El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido

tomada de otros documentos está debidamente referenciada.

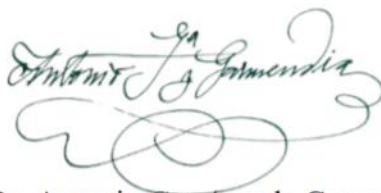
A handwritten signature in black ink, appearing to read 'Eduardo Ortega Isasi', with a large, stylized flourish extending from the bottom left.

Fdo.: Eduardo Ortega Isasi

Fecha: 29 de JUNIO de 2021

Autorizada la entrega del proyecto

EL DIRECTOR DEL PROYECTO

A handwritten signature in black ink, appearing to read 'Antonio García y de Garmendia', with a large, stylized flourish extending from the bottom left.

Fdo.: Dr. Antonio García y de Garmendia

Fecha: 28 de JUNIO de 2021

AUTORIZACIÓN PARA LA DIGITALIZACIÓN, DEPÓSITO Y DIVULGACIÓN EN RED DE PROYECTOS FIN DE GRADO, FIN DE MÁSTER, TESINAS O MEMORIAS DE BACHILLERATO

1º. Declaración de la autoría y acreditación de la misma.

El autor D. EDUARDO ORTEGA ISASI DECLARA ser el titular de los derechos de propiedad intelectual de la obra: MODELO PREDICTIVO DE INVERSIÓN EN OBRAS DE ARTE, que ésta es una obra original, y que ostenta la condición de autor en el sentido que otorga la Ley de Propiedad Intelectual.

2º. Objeto y fines de la cesión.

Con el fin de dar la máxima difusión a la obra citada a través del Repositorio institucional de la Universidad, el autor **CEDE** a la Universidad Pontificia Comillas, de forma gratuita y no exclusiva, por el máximo plazo legal y con ámbito universal, los derechos de digitalización, de archivo, de reproducción, de distribución y de comunicación pública, incluido el derecho de puesta a disposición electrónica, tal y como se describen en la Ley de Propiedad Intelectual. El derecho de transformación se cede a los únicos efectos de lo dispuesto en la letra a) del apartado siguiente.

3º. Condiciones de la cesión y acceso

Sin perjuicio de la titularidad de la obra, que sigue correspondiendo a su autor, la cesión de derechos contemplada en esta licencia habilita para:

- a) Transformarla con el fin de adaptarla a cualquier tecnología que permita incorporarla a internet y hacerla accesible; incorporar metadatos para realizar el registro de la obra e incorporar “marcas de agua” o cualquier otro sistema de seguridad o de protección.
- b) Reproducir la en un soporte digital para su incorporación a una base de datos electrónica, incluyendo el derecho de reproducir y almacenar la obra en servidores, a los efectos de garantizar su seguridad, conservación y preservar el formato.
- c) Comunicarla, por defecto, a través de un archivo institucional abierto, accesible de modo libre y gratuito a través de internet.
- d) Cualquier otra forma de acceso (restringido, embargado, cerrado) deberá solicitarse expresamente y obedecer a causas justificadas.
- e) Asignar por defecto a estos trabajos una licencia Creative Commons.
- f) Asignar por defecto a estos trabajos un HANDLE (URL *persistente*).

4º. Derechos del autor.

El autor, en tanto que titular de una obra tiene derecho a:

- a) Que la Universidad identifique claramente su nombre como autor de la misma
- b) Comunicar y dar publicidad a la obra en la versión que ceda y en otras posteriores a través de cualquier medio.
- c) Solicitar la retirada de la obra del repositorio por causa justificada.
- d) Recibir notificación fehaciente de cualquier reclamación que puedan formular terceras personas en relación con la obra y, en particular, de reclamaciones relativas a los derechos de propiedad intelectual sobre ella.

5º. Deberes del autor.

El autor se compromete a:

- a) Garantizar que el compromiso que adquiere mediante el presente escrito no infringe ningún derecho de terceros, ya sean de propiedad industrial, intelectual o cualquier otro.
- b) Garantizar que el contenido de las obras no atenta contra los derechos al honor, a la intimidad y a la imagen de terceros.
- c) Asumir toda reclamación o responsabilidad, incluyendo las indemnizaciones por daños, que

d) Asumir la responsabilidad en el caso de que las instituciones fueran condenadas por infracción de derechos derivada de las obras objeto de la cesión.

La obra se pondrá a disposición de los usuarios para que hagan de ella un uso justo y respetuoso con los derechos del autor, según lo permitido por la legislación aplicable, y con fines de estudio, investigación, o cualquier otro fin lícito. Con dicha finalidad, la Universidad asume los siguientes deberes y se reserva las siguientes facultades:

- Madrid, a 29 de JUNIO de 2021

Fdo Eduardo Ortega Isasi

Edward S. Rogers

Motivos para solicitar el acceso restringido, cerrado o embargado del trabajo en el Repositorio Institucional:





GRADO EN INGENIERÍA EN TECNOLOGÍAS INDUSTRIALES

TRABAJO FIN DE GRADO

MODELO PREDICTIVO DE INVERSIÓN EN OBRAS DE ARTE

Ortega Isasi, Eduardo

Director: Dr. Antonio García y de Garmendia

Madrid

Julio 2021



Escuela Técnica Superior de Ingeniería (ICAI)
Ortega Isasi, Eduardo
Modelo Predictivo de Inversión en Obras de Arte

MODELO PREDICTIVO DE INVERSIÓN EN OBRAS DE ARTE

Ortega Isasi, Eduardo

Director: Dr. García y de Garmendia, Antonio

Universidad Pontificia Comillas

RESUMEN DEL PROYECTO

El proyecto consiste en el desarrollo de un modelo de predicción de obras de arte de autores desconocidos basado en una regresión hedónica combinándolo con el precio de venta anterior en aquellas obras en las que se disponga de dicha información. Se han identificado relaciones significativas entre el precio de venta, el siglo de creación, el origen de la obra, la casa de subastas en la que se vende y la aparente suavidad de la obra.

Palabras clave: Arte, Pintura, Predicción, Precio, Inversión

1. Introducción

La industria del arte está muy enfocada a la especulación con obras artísticas en base a pintores con gran renombre. No obstante, el día a día de esta industria responde a otro tipo de obras de artistas desconocidos en las que predominan los atributos visibles de la obra para su valoración.

2. Definición del Proyecto

El proyecto busca continuar incrementando la información y la investigación enfocada a la consideración de la industria del arte como una de las nuevas inversiones alternativas. El estado del arte nos desvela que toda la investigación desarrollada hasta la actualidad se encuentra enfocada en obras de alto valor realizadas por autores con un gran renombre.

Por ello, este proyecto se enfoca en el sector contrario de la industria del arte, es decir, aquellas obras de un menor valor y que no disponen de un autor reconocido o su autor no tiene gran influencia en el precio de venta. El precio de este tipo de

obras suele tender en gran medida a determinarse por las características físicas y visuales de la obra y el gusto del comprador, a diferencia del arte especulativo donde el valor de la obra reside en la firma del autor en cuestión.

3. Descripción del modelo

El modelo de predicción descrito anteriormente consta de dos partes:

- Una regresión hedónica en la que la estimación del precio de venta de la obra se genera a partir de un conjunto de variables visibles de dicha obra. La estimación de los parámetros de la regresión se calcula mediante la metodología de Mínimos Cuadrados Ordinales (MCO).
- En aquellos casos en los que se dispone de la información del precio de venta de una obra en un momento temporal pasado, esta información se combina con el modelo anterior.
- Cuando se dispone de estas dos predicciones, se combinan de manera proporcional con el objetivo de obtener una predicción final. Como se podrá comprobar, la estimación calculada con este modelo combinado es más precisa que la calculada únicamente a partir de la regresión hedónica.

La metodología empleada para el desarrollo del modelo se representa a continuación en un diagrama de flujo.

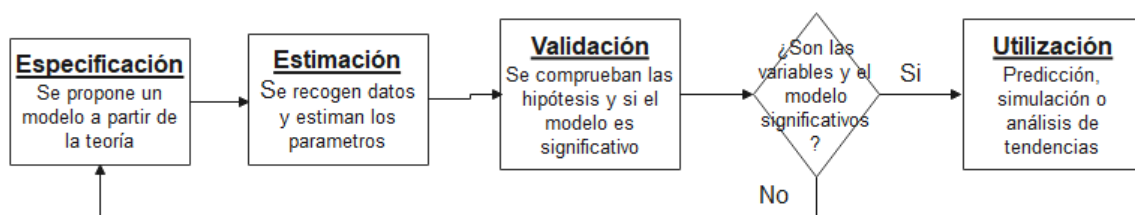


Ilustración 1 – Metodología para el desarrollo del modelo predictivo. Fuente: Elaboración Propia (2021)

4. Resultados

El estudio concluye con la elaboración del modelo predictivo para la valoración de obras de autores desconocidos descrito en el apartado anterior y la identificación de las siguientes tendencias entre las variables analizadas:

- Una relación negativa entre el siglo de creación de la obra y su precio de venta, es decir, las obras de mayor antigüedad tienden a venderse por un precio mayor.

- Las obras artísticas realizadas en América muestran menores precios que aquellas realizadas en Europa y Asia con un nivel de significación superior al 90% y 99% respectivamente.
- Las obras que se alejan de los vértices y esquinas y buscan la suavidad en sus formas tienden a venderse por un precio mayor.
- La casa de subastas Clars Auction Gallery tienen un precio de venta de las obras mayor en comparación a otras casas de subastas como Jasper 52 o Hudson Valley Auctions.

A continuación se presenta una grafica del modelo puesto en práctica. En la gráfica se representan las predicciones del modelo para las observaciones 15-50, un intervalo de confianza y el precio real de venta.

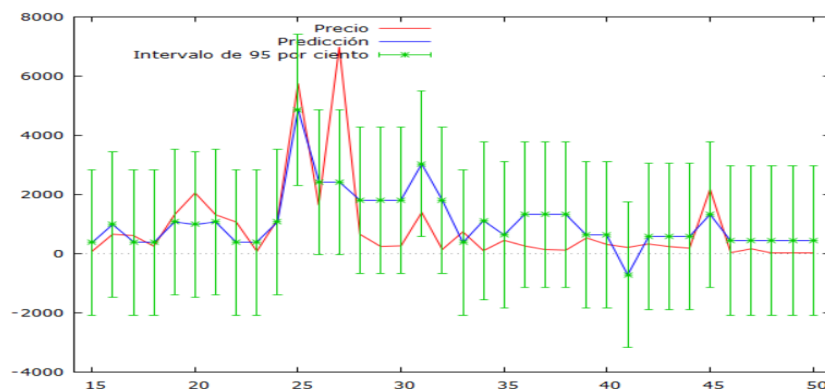


Ilustración 2 – Predicción vs Precio real de la observación 15 – 50. Fuente: Elaboración Propia (2021)

5. Conclusiones

Las conclusiones más destacadas de este estudio se presentan a continuación:

- La aplicación de técnicas predictivas a la industria del arte no está generalizada, no obstante, puede traer multitud de beneficios para todos los participantes en este sector mediante la reducción de costes, tiempo y esfuerzo, junto con el aumento de la precisión.
- La combinación de nuestro modelo predictivo con el conocimiento de los expertos tasadores de las casas de subastas puede permitirles partir de una predicción inicial y realizar ajustes en aquellas ocasiones en las que el modelo no pueda captar nuevas tendencias o influencias en el precio.

- La predicción del valor de obras de arte mediante modelos basados en regresiones hedónicas aporta numerosos beneficios dada la transparencia que ofrece esta técnica la cual permite conocer el valor de todos los parámetros y variables para su modificación cuando es necesario.

6. Referencias

Abidoye, R. B., & Chan, A. P. (Febrero de 2018). *Improving property valuation accuracy: A comparison of hedonic pricing model and artificial neural network*. Pacific Rim Property Research Journal.

Arte, Capital del (10 de Enero de 2018). *Arte figurativo – Características*.

Aubry, M., Kräussl, R., Manso, G., & Spaenjers, C. (Julio de 2020). *Biased Auctioneers*. HEC Paris.

Bellver, J. A., & Martínez, F. G. (2007). *Modelos de valoración económica de obras de arte pictóricas*. Dialnet.

Blasco, I. M. (18 de Febrero de 1998). *Modelos econométricos aplicados a la valoración de bienes inmuebles rústicos*. Dialnet.

Burke, J., & Burns, E. (Marzo de 2021). *What is a neural network? Explanation and examples*. TechTarget.

Fernando, J., & Khartit, K. (13 de Marzo de 2021). *Discounted Cash Flow (DCF)*. Investopedia.

Guillén, A. M. (2015). *Modelo de estimación del precio de las inversiones en obras de arte a través de redes neuronales: La pintura de Picasso en el mercado de las subastas*. Universidad Rey Juan Carlos.

Ghosh, Sourabh, & Mehta, Kiran (2017). *Art as an Investment: Factors Affecting Valuation of Miniature Paintings in India*. Department of Management Studies, IIT Roorkee.

Hodgson, D., & Galbraith, J. (Junio de 2018). *Econometric Fine Art Valuation by Combining Hedonic and Repeat-Sales Information*. Research Gate.

Hosny, A. (9 de Diciembre de 2017). *Machine Learning For Art Valuation. An Interview With Ahmed Hosny*. Artnome.

Koehrsen, W. (28 de Enero de 2018). *Overttting vs. Underttting: A*

Complete Example. Patuakhali Science and Technology University.

Michelone, & López, M. (9 de Julio de 2019). *¿Cómo funcionan las redes neuronales?*. Unocero.

Rodríguez, L. (2021). *Qué son las redes neuronales y cuál es su aplicación en el marketing.* Artyco.

Tabales, J. M., Carmona, F. J., & Ocerin, J. M. (29 de Marzo de 2017). *Neural networks (NN) applied to the commercial properties valuation.* Informes de la Construcción.

Trigo, A (2021). *Criterios para Valorar una Obra de Arte.*

Zorita, C. M., Álvarez, C., Borrás, F., & Budría, S. (2019). *Modelos Cuantitativos para la Economía y la Empresa en 101 ejemplos.* EV Services.

Carollo, M. Carmen (2011-2012). *Contrastes de Hipótesis.* Universidad de Santiago de Compostela

MC Larrosa, Juan (2021). *Econometría: Regresión Hedónica.*

MODELO PREDICTIVO DE INVERSIÓN EN OBRAS DE ARTE

Ortega Isasi, Eduardo

Supervisor: Dr. García y de Garmendia, Antonio

Universidad Pontificia Comillas

ABSTRACT

The project aims to develop a predictive model for artwork from unknown artists based in a hedonic regression combined with the prices of previously sold artwork in those cases when we have access to that information. We have identified significant relationships between the following variables: selling price, century, origin, auction house and smoothness.

Keywords: Art, Painting, Prediction, Price, Investing

1. Introduction

The art industry is primarily focused on speculation with artwork from famous artists. Nevertheless, the day-to-day work of this industry is mainly based on artwork from unknown artists which its selling price is driven by their visible characteristics.

2. Definition of the project

The project aims to increase the available information and continue to explore the art industry considering it as one of the new areas of alternative investing. The state-of-the-art reveals that all the investigation carried out is focused on analyzing artwork from famous artists.

Those reasons drive us to focus our project on other direction in which we analyze the art industry from a different point of view. We analyze lower valued pieces of art in which visible characteristics determine its selling price instead of being driven by its artists' name.

3. Description of the model

The previously discussed model consists of two different parts:

- A hedonic regression in which an estimation of the selling price is computed from a set of visible variables of that piece of art. The estimation of the regression parameters is computed following the Ordinary Least Square (OLS) methodology.
- In those cases in which we have the information of the selling price of that previously sold piece of art, this information is combined with the previously described model.
- When we have those two estimations, they are combined in a proportional way aiming to obtain a final prediction. As we have demonstrated, the combined estimation of the selling price is more precise than that computed only with the hedonic regression model.

The methodology used for the development of the model is presented in the following flowchart.

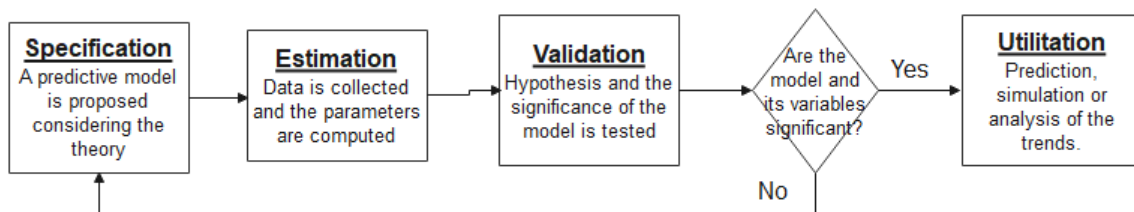


Figure 1. Methodology for the development of the predictive model. Source: Prepared by the authors (2021)

4. Results

The project concludes with the development of the predictive model for unknown artists previously described and the identification of the following trends between the variables analyzed:

- Negative relationship between the creation century and the selling price of artwork. Older pieces of art tend to be sold for higher prices.
- Those pieces of art made in America tend to be sold for lower prices than those pieces of art which are made in Europe or Asia, with a significance level of 90% and 99% respectively.

- Those pieces of art which are made in a smoother style walking away from corners and edges tend to be sold for a higher price.
- Artwork sold in Clars Auction Gallery have a higher selling price compared with artwork sold in Jasper 52 or Hudson Valley Auctions.

In the following graph, a practical example of the final model is presented. In this graph we can see the observed and predictive selling prices and a confidence interval of that prediction.

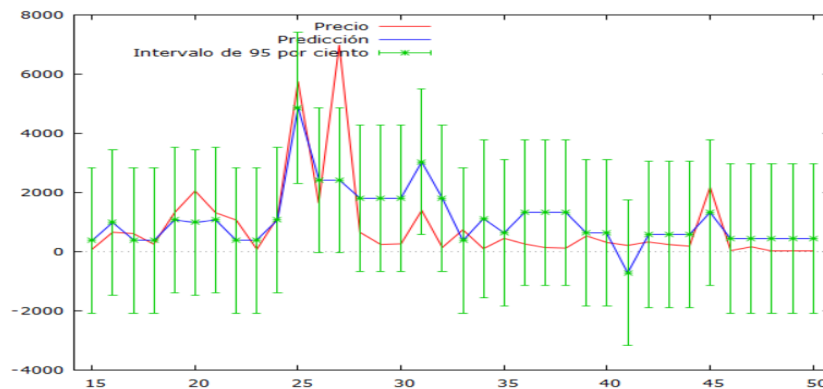


Figure 2. Prediction vs Observed Price from observation 15 to 50. Source: Prepared by the authors (2021)

5. Conclusions

The key conclusions were as follows:

- The application of predictive techniques to the art industry is not widespread, nevertheless, it can bring multiple benefits to all the participants in the industry through cost, time and effort reduction and a increase in the precision of the estimations.
- The combination of our predictive model with the knowledge of an expert appraiser can let them start their predictions from an initial selling price and then, make any adjustments needed when the hedonic model cannot capture new trends or influences in the selling price.
- The prediction of the selling price of artwork through hedonic regression models brings large advantages because of the transparency of this technique. It enables its users to know the value of all the parameters and variables and those expert users can change the values when needed.

6. References

- Abidoye, R. B., & Chan, A. P. (Febrero de 2018). *Improving property valuation accuracy: A comparison of hedonic pricing model and artificial neural network*. Pacific Rim Property Research Journal.
- Arte, Capital del (10 de Enero de 2018). *Arte figurativo – Características*.
- Aubry, M., Kräussl, R., Manso, G., & Spaenjers, C. (Julio de 2020). *Biased Auctioneers*. HEC Paris.
- Bellver, J. A., & Martínez, F. G. (2007). *Modelos de valoración económica de obras de arte pictóricas*. Dialnet.
- Blasco, I. M. (18 de Febrero de 1998). *Modelos econométricos aplicados a la valoración de bienes inmuebles rústicos*. Dialnet.
- Burke, J., & Burns, E. (Marzo de 2021). *What is a neural network? Explanation and examples*. TechTarget.
- Fernando, J., & Khartit, K. (13 de Marzo de 2021). *Discounted Cash Flow (DCF)*. Investopedia.
- Guillén, A. M. (2015). *Modelo de estimación del precio de las inversiones en obras de arte a través de redes neuronales: La pintura de Picasso en el mercado de las subastas*. Universidad Rey Juan Carlos.
- Ghosh, Sourabh, & Mehta, Kiran (2017). *Art as an Investment: Factors Affecting Valuation of Miniature Paintings in India*. Department of Management Studies, IIT Roorkee.
- Hodgson, D., & Galbraith, J. (Junio de 2018). *Econometric Fine Art Valuation by Combining Hedonic and Repeat-Sales Information*. Research Gate.
- Hosny, A. (9 de Diciembre de 2017). *Machine Learning For Art Valuation. An Interview With Ahmed Hosny*. Artnome.
- Koehrsen, W. (28 de Enero de 2018). *Overttting vs. Underttting: A Complete Example*. Patuakhali Science and Technology University.
- Michelone, & López, M. (9 de Julio de 2019). *¿Cómo funcionan las redes neuronales?*. Unocero.
- Rodríguez, L. (2021). *Qué son las redes neuronales y cuál es su aplicación en el marketing*. Artyco.

Tabales, J. M., Carmona, F. J., & Ocerin, J. M. (29 de Marzo de 2017). *Neural networks (NN) applied to the commercial properties valuation*. Informes de la Construcción.

Trigo, A (2021). *Criterios para Valorar una Obra de Arte*.

Zorita, C. M., Álvarez, C., Borrás, F., & Budría, S. (2019). *Modelos Cuantitativos para la Economía y la Empresa en 101 ejemplos*. EV Services.

Carollo, M. Carmen (2011-2012). *Contrastes de Hipótesis*. Universidad de Santiago de Compostela

MC Larrosa, Juan (2021). *Econometría: Regresión Hedónica*.



Escuela Técnica Superior de Ingeniería (ICAI)
Ortega Isasi, Eduardo
Modelo Predictivo de Inversión en Obras de Arte

ÍNDICE GENERAL

1.	INTRODUCCIÓN.....	1
1.1.	Los modelos predictivos y la industria del arte	1
1.2.	Objetivos y resultados esperados.....	3
1.3.	Motivaciones	3
1.4.	Metodología y estructura del proyecto	5
2.	ESTADO DEL ARTE	6
3.	DISEÑO DE UN MODELO DE PREDICCIÓN DE OBRAS DE ARTE	18
3.1.	Justificación de la selección del área de desarrollo del modelo predictivo	18
3.2.	Selección de las técnicas del modelo predictivo	20
3.3.	Selección del conjunto de variables explicativas	29
3.4.	Descripción estadística de la base de datos empleada para la elaboración del modelo predictivo.....	38
3.5.	Especificación del modelo de predicción	40
3.6.	Conclusiones.....	45
4.	VALIDACIÓN DEL MODELO DE PREDICCIÓN.....	46
4.1.	Introducción.....	46
4.2.	Estimación y análisis de los modelos de predicción.....	46
4.3.	Estimadores de los parámetros de la regresión.....	60
4.4.	Inclusión de la información de ventas pasadas en el modelo	63
4.5.	Predicción con el modelo predictivo	67
4.6.	Conclusiones.....	71
5.	RESULTADOS, CONCLUSIONES Y DESARROLLOS FUTUROS	73
5.1.	Introducción.....	73
5.2.	Resultados, conclusiones y comparación con otros estudios	73
5.3.	Desarrollos Futuros	79
6.	BIBLIOGRAFÍA	82
7.	ANEXOS.....	84

ÍNDICE DE ILUSTRACIONES

Ilustración 1 – Metodología para el desarrollo del modelo predictivo. Fuente: Elaboración Propia (2021).....	11
Ilustración 2 – Predicción vs Precio real de la observación 15 – 50. Fuente: Elaboración Propia (2021).....	12
Ilustración 3. Fuente: Elaboración Propia (2021).....	22
Ilustración 4. Fuente: Elaboración Propia (2021).....	23
Ilustración 5. Fuente: Elaboración Propia (2021).....	25
Ilustración 6. Fuente: Elaboración Propia (2021).....	27
Ilustración 7. Fuente: Elaboración Propia (2021).....	30
Ilustración 8. Fuente: Elaboración Propia (2021).....	31
Ilustración 9. Fuente: Elaboración Propia (2021).....	32
Ilustración 10. Obra en un color. Fuente: Elaboración Propia (2021).....	33
Ilustración 11. Obra en varios colores. Fuente: Elaboración Propia (2021).....	33
Ilustración 12. Obra con baja nitidez. Fuente: Elaboración Propia (2021)	34
Ilustración 13. Obra con alta nitidez. Fuente: Elaboración Propia (2021)	34
Ilustración 14. Obra no suave. Fuente: Elaboración Propia (2021)	34
Ilustración 15. Obra suave. Fuente: Elaboración Propia (2021).....	34
Ilustración 16. Paisaje. Fuente: Elaboración Propia (2021)	35
Ilustración 17. Retrato/Escena. Fuente: Elaboración Propia (2021)	35
Ilustración 18. Estilo surrealista. Fuente: Elaboración Propia (2021).....	36
Ilustración 19. Estilo abstracto. Fuente: Elaboración Propia (2021).....	36
Ilustración 20. Obra clara. Fuente: Elaboración Propia (2021).....	37
Ilustración 21. Obra oscura. Fuente: Elaboración Propia (2021)	37
Ilustración 22. Fuente: Elaboración Propia (2021).....	39
Ilustración 23. Fuente: Elaboración Propia (2021).....	40
Ilustración 24. Fuente: Elaboración Propia (2021).....	44

Ilustración 25. Fuente: Álvarez et al. (2021).....	48
Ilustración 26. Fuente: Álvarez et al. (2021).....	50
Ilustración 27. Fuente: Elaboración Propia	54
Ilustración 28. Fuente: Elaboración Propia (2021).....	56
Ilustración 29. Fuente: Elaboración Propia (2021).....	58
Ilustración 30. Fuente: Elaboración Propia (2021).....	59
Ilustración 31. Fuente: Elaboración Propia (2021).....	60
Ilustración 32. Fuente: Elaboración Propia (2021).....	61
Ilustración 33. Fuente: Elaboración Propia (2021).....	62
Ilustración 34. Fuente: Elaboración Propia (2021).....	64
Ilustración 35. Fuente: Elaboración Propia (2021).....	65
Ilustración 36. Fuente: Elaboración Propia (2021).....	68
Ilustración 37. Fuente: Elaboración Propia (2021).....	69
Ilustración 38. Fuente: Elaboración Propia (2021).....	81
Ilustración 39. Fuente: Elaboración Propia (2021).....	88
Ilustración 40. Fuente: Elaboración Propia (2021).....	89

ÍNDICE DE TABLAS

Tabla 1. Fuente: Elaboración Propia (2021).....	41
Tabla 2. Fuente: Elaboración Propia (2021).....	43
Tabla 3. Fuente: Elaboración Propia (2021).....	64
Tabla 4. Fuente: Elaboración Propia (2021).....	66
Tabla 5. Fuente: Elaboración Propia (2021).....	70
Tabla 6. Fuente: Elaboración Propia (2021).....	71
Tabla 7. Fuente: Elaboración Propia (2021).....	85
Tabla 8. Fuente: Elaboración Propia (2021).....	86

1. INTRODUCCIÓN

1.1. Los modelos predictivos y la industria del arte

El mundo se ha visto inmerso en una rápida evolución de la tecnología en el último siglo. La sociedad ha evolucionado de la mano de este desarrollo tecnológico y se ha adaptado progresivamente, lo que ha cambiado la manera de percibir y realizar la gran parte de las actividades de nuestra sociedad. La tecnología ha traído innumerables mejoras para la sociedad de la actualidad, facilitando la realización de numerosas actividades, incrementando el acceso a información gratuita y de calidad o globalizando e interconectando las distintas regiones del planeta. Las empresas y corporaciones del mundo han aprovechado la oportunidad que ofrece esta globalización y digitalización para mejorar sus actividades, realizándolas de forma más eficiente. Ello ha generado un auge de la competitividad que ha llevado a que sea de vital importancia la disposición de sistemas o modelos de predicción que les permitan disfrutar de estas ventajas.

En las últimas dos décadas, la disposición y recolección de información ha generado grandes bases de datos que suponen una gran fuente de poder para gobiernos, empresas o individuos. Ello ha facilitado y resaltado aún más la utilización de estas técnicas de predicción que se basan en la disponibilidad de gran cantidad de datos. Las empresas e instituciones están invirtiendo más que nunca en nuevas tecnologías como el “machine learning” así como profesionales expertos en “big data” que les permitan gestionar y sacarles partido a estas grandes bases de datos, con el fin de analizarlas y desarrollar modelos predictivos con usos tan variados como la predicción de la demanda o los precios de venta en subastas.

En el sector de la inversión, los métodos de análisis clásicos empleados durante muchos años han dejado paso a métodos cuantitativos de predicción que trabajan con gran cantidad de datos y variables. Gracias a la tecnología actual, estos métodos nos permiten obtener valoraciones y conclusiones más precisas, en menor tiempo y con un limitado margen de error. Todo ello desde un punto de vista mucho más objetivo.

Los modelos predictivos han evolucionado como la principal herramienta para valorar cualquier tipo de activo, desde el mercado inmobiliario, pasando por participaciones de compañías hasta objetos de colección. Gracias a ello, se pueden conseguir estimaciones precisas a un coste reducido frente al que conllevaría la utilización de otras técnicas que tienen mayor error o la consulta de un experto en la materia, lo que aumenta considerablemente el coste de la estimación. También permite a los expertos ahorrar trabajo, ya que pueden desarrollar sus análisis teniendo en cuenta las estimaciones generadas por modelos de predicción.

El arte siempre ha sido uno de los pilares de la decoración. No obstante, en las últimas décadas ha empezado a considerarse como una vía de inversión alternativa. No obstante, cabe diferenciar dos lecturas distintas de la industria del arte. Por un lado, encontramos el arte especulativo, aquel en el que el nombre del autor marca la diferencia por su fama o renombre, lo que nos lleva a obras artísticas que se disparan en su precio simplemente por ser realizadas por cierto autor. Por otro lado, el arte más puro, aquel en el que el nombre del autor no juega un papel importante y son las características de la propia obra las que nos llevan a una determinada valoración. En próximos capítulos diferenciaremos estas dos variantes del arte en cuanto a sus diferencias en la especificación del modelo predictivo.

1.2. **Objetivos y resultados esperados**

El objetivo principal del proyecto es el desarrollo de un modelo de predicción para la valoración de obras de arte, junto con la identificación de tendencias entre las distintas variables que analizaremos relacionadas con la industria del arte. Todo ello con el objetivo de poder seguir aumentando la información disponible a cerca del análisis de esta industria desde la perspectiva de la valoración e inversión. También se busca la elaboración de un modelo que facilite el trabajo de tasación de las casas de subastas.

Para ello, analizaremos las distintas tecnologías predictivas existentes en la actualidad de cara a seleccionar una o varias de ellas como la base de nuestro modelo. Posteriormente, analizaremos como afectan diferentes variables características de la industria del arte en la determinación del precio de venta de diversas obras artísticas. Todo ello partiendo de la generación manual de una base de datos con aquellas variables cuyo estudio sea de interés.

A lo largo del estudio se han desvelado diversas relaciones sorprendentes entre variables, las cuales pueden ser diferentes a las ideas preconcebidas que puedan tener los lectores, que nos lleven a la visión de la industria del arte desde una nueva perspectiva tras la lectura del proyecto.

1.3. **Motivaciones**

Una vez explicado el contexto en el que se desarrolla el proyecto, continuamos con las razones que me llevan a la elaboración del mismo. Como se ha comentado anteriormente, el auge en el análisis “big data” y los modelos predictivos por todos los beneficios que estas técnicas nos ofrecen junto con la creciente visión del arte como una

importante vía de inversión alternativa en los últimos años, me llevan a desarrollar un proyecto de investigación que fusione ambos aspectos y trate de estudiar esta industria desde una perspectiva novedosa.

Adicionalmente, otra razón fundamental para el desarrollo del proyecto es el escaso progreso e información que existe en las técnicas de valoración de obras artísticas de autores desconocidos en comparación con la información relevante al estudio de obras de autores famosos o la investigación disponible en otras industrias como los mercados financieros o los bienes inmuebles en las que estas técnicas de predicción están ampliamente extendidas. Por ello considero que existen carencias en este ámbito, lo que me llevaba a realizar un proyecto que trate de evaluar las diferentes técnicas predictivas con el objetivo de determinar cuál o cuáles son las más efectivas para la valoración del arte de autores con un menor renombre.

Finalmente, las industrias mencionadas anteriormente en las que se han aplicado técnicas de modelaje predictivo están dominadas por activos que se caracterizan por variables objetivas. No obstante, la industria del arte está dominada por multitud de variables subjetivas, lo que supone un interesante reto a la hora de representarlas en un modelo de predicción y comprobar si realmente es factible introducir estas técnicas en esta industria.

Por todo lo mencionado anteriormente, este proyecto supone una enorme oportunidad y experiencia de investigación en un sector que está mostrando su capacidad como una de las grandes inversiones alternativas disponibles en la actualidad y una técnica que ya lleva años creciendo y estableciéndose como una de las más importantes.

1.4. **Metodología y estructura del proyecto**

El proyecto se estructura como se muestra a continuación:

- En primer lugar, se procede con el análisis del estado del arte de cara a estudiar la información de las distintas técnicas de predicción existentes en la actualidad y su aplicación para la valoración de activos en la industria del arte y en otras industrias más investigadas.
- Posteriormente, se procede con el estudio de las ventajas e inconvenientes de las diversas técnicas aplicadas a la valoración de obras de arte de autores desconocidos, junto con la elaboración final del modelo mediante la selección de las variables consideradas relevantes. También, se recogerá de manera manual una base de datos con dichas variables seleccionadas.
- A continuación se procede con la estimación de los parámetros y el análisis de las variables de cara a determinar aquellas relevantes para la definición del modelo final de predicción. De manera adicional, buscamos la identificación de diversas tendencias entre las variables analizadas de la industria del arte.
- Finalmente, se presentarán los resultados y conclusiones obtenidos en el proyecto junto con la comparación con otros estudios que hayan estudiado modelo y variables similares.

2. ESTADO DEL ARTE

Como se ha comentado en la introducción, el análisis predictivo y los modelos de valoración han ganado mucha fuerza en el último siglo con el auge de la tecnología y la disponibilidad de información a través de internet, junto con la aparición de herramientas como el “machine learning” o el “data mining” que permiten analizar bases de datos muy amplias y complejas en un reducido intervalo de tiempo. Estos avances tecnológicos son aplicados en multitud de áreas desde la identificación del público objetivo para una empresa hasta la predicción de la demanda que un negocio tendrá en los próximos meses. No obstante, nos centraremos en el estado del arte de las técnicas de análisis predictivo para la valoración de activos.

Comenzaremos analizando las diversas técnicas predictivas existentes hasta la actualidad para la valoración de activos de los que se dispone de mayor cantidad de información, como pueden ser los mercados financieros o los bienes inmuebles. Junto con el análisis de cada técnica, mencionaremos aquellos proyectos e investigaciones que han centrado su estudio en el estado del arte y se han enfocado en las variables representativas o atributos de las obras artísticas con el objetivo de predecir el valor de este tipo de activos.

En primer lugar, cuando un activo genera flujos monetarios como una participación en una compañía que reparte dividendos o un proyecto del que se espera obtener ingresos en un futuro, una de las técnicas¹ más utilizadas consiste en descontar los flujos esperados en un futuro y sumar dichas cuantías en el presente para calcular su valor

¹ Fernando & Khartit (2021)

actual, siendo el activo con mayor valoración aquel que ofrezca una mayor suma en el presente. Esta técnica² es una de la más utilizada a la hora de valorar las participaciones de corporaciones que reparten dividendos. No obstante, por norma general, las obras de arte no generan flujos monetarios para sus propietarios por lo que descartaremos seguir ahondando en esta técnica.

A continuación, se presentan los avances existentes hasta la actualidad en una de las técnicas que ha experimentado un mayor incremento en su uso y popularidad en los últimos años, conocida como las redes neuronales.

Las redes neuronales son simulaciones en software³ que tratan de imitar el funcionamiento de un cerebro humano uniendo millones de neuronas artificiales⁴ entre sí con el objetivo de que pueda aprender, reconocer patrones e incluso tomar decisiones. El sistema se alimenta con una gran cantidad de datos de entrada y se compara la salida del software con el resultado correcto, registrando el error y tratando de reducirlo de manera progresiva. Esta dinámica permite que la red aprenda a predecir el valor de salida esperado. Una vez el sistema ha sido entrenado correctamente podemos comenzar a emplearlo⁵ con nuevos datos para predecir el valor de aquello para lo que ha sido entrenando en un primer lugar.

² Conocida como Flujo de Fondos Descontados.

³ López & Michelone (2019)

⁴ Marcias Guillén (2015). Cada neurona artificial realiza una función matemática y esta interconectada con las demás neuronas artificiales mediante enlaces de comunicación.

⁵ López & Michelone (2019)

Las redes neuronales han ganado mucha fuerza en los últimos años en el área de la predicción del valor de activos ya que ofrecen multitud de ventajas⁶ frente a otras técnicas como el análisis predictivo de regresión o la valoración por comparación.

En líneas generales, los modelos basados en redes neuronales ofrecen resultados más precisos y con menor error en la gran mayoría de sus aplicaciones. En concreto, como se puede comprobar en los estudios que se mencionan a continuación, una de sus cualidades principales es que dichos modelos predictivos basados en redes neuronales artificiales tienen una mejor capacidad para reflejar las relaciones no lineales y adaptarse a bases de datos que presentan resultados extremos “outliers”. Una comparación entre un modelo⁷ basado en una red neuronal frente a otro que consistía en una regresión hedónica múltiple, ambos con el objetivo de predecir el valor de locales comerciales⁸, mostró como el modelo basado en una red neuronal superó en términos de precisión al modelo basado en una regresión hedónica con un grado de ajuste muy superior⁹, mayor correlación¹⁰ entre observaciones y estimaciones, así como errores cometidos muy inferiores al modelo de regresión.

Otro estudio¹¹ muy similar al mencionado anteriormente comparó¹² los mismos dos modelos con el fin de predecir el valor de bienes inmuebles residenciales. Los resultados fueron muy similares al estudio comentado en el párrafo anterior, con unos errores medios de porcentaje absoluto¹³ de 15.94% (modelo basado en una red neuronal) y 38.23% (modelo basado en una regresión hedónica). En ambos estudios

⁶ En ciertas aplicaciones, las características de las redes neuronales provocan que no sea la técnica más adecuada. En el siguiente capítulo se profundizará en este aspecto.

⁷ Núñez, Rey & Caridad y Ocerin (2017)

⁸ Situados en la ciudad española de Córdoba. Este dato será relevante ya que centrar el modelo en un área concreta tiene beneficios en la precisión del mismo como veremos más adelante.

⁹ Coeficiente de determinación (R^2) de 87,38 % frente al 71,60 % respectivamente. Este coeficiente determina la calidad del modelo para replicar los resultados.

¹⁰ Coeficiente de correlación de 0,8396 a 0,9352 respectivamente. Este coeficiente es una medida de la dependencia lineal entre variables.

¹¹ Publicado en 2018 en el “*Pacific Rim Property Research Journal*”

¹² Boluwatife & Chan (2018)

¹³ Medidos con el indicador estadístico “Mean Absolute Percentage Error” (MAPE).

podemos comprobar como la diferencia entre ambos modelos se origina por las características del mercado al que están tratando de ajustarse. El modelo basado en una regresión hedónica asume que las variaciones¹⁴ en el precio estimado del inmueble son siempre lineales frente a variaciones en las variables explicativas, situación que no se corresponde con la realidad del mercado y que la red neuronal corrige de una forma más congruente por su superioridad en la representación de relaciones no lineales como se ha comentado anteriormente.

No obstante, cabe destacar que los modelos basados en redes neuronales también presentan una serie de desventajas¹⁵ entre las que destacan un tiempo de aprendizaje elevado que depende de la cantidad de patrones a identificar y la flexibilidad de la red neuronal para reconocer patrones que sean sumamente parecidos, junto con la necesidad de una elevada cantidad de datos para el entrenamiento y la necesidad de disponer de un operario que interprete la salida proporcionada por la red.

Las redes neuronales también han sido aplicadas de manera exitosa en la predicción del valor de diferentes activos financieros operados en los mercados bursátiles, pero el mercado del arte presenta un menor número de proyectos e investigaciones en los que se haya usado esta técnica de predicción. No obstante, el estudio en profundidad de la aplicación¹⁶ de las redes neuronales como modelo de predicción del valor de las pinturas del artista Pablo Picasso, permitió obtener un modelo con un coeficiente de correlación de 98.23% y un error absoluto medio de 23.35%.

¹⁴ Núñez, Rey & Caridad y Ocerin (2017)

¹⁵ Rodríguez (2021)

¹⁶ Marcías Guillén (2015)

La mayor parte de las obras de Pablo Picasso tuvieron un precio de remate superior a los 500.000 €. No obstante, 10 de las 395 obras analizadas tuvieron un precio de remate inferior a los 8.000€. La predicción del valor de estas obras no fue nada precisa, lo cual nos muestra que la presunta superioridad en la predicción de valores extremos *outliers* de los modelos basados en redes neuronales frente a otro tipo de modelos no siempre se mantiene y depende en gran medida de las características del mercado al que se ajusta el modelo.

También es importante destacar que en dicho estudio¹⁷ se propusieron 11 variables consideradas relevantes en la predicción del valor de las pinturas de Pablo Picasso. Finalmente, se concluyó que la combinación de 6 de dichas variables generaba el menor error de predicción posible. Las variables seleccionadas fueron el precio de salida, el año de venta, el país de venta, la casa de subastas, la bibliométrica y la redmétrica.

Otra de las técnicas más empleadas en la actualidad a la hora de realizar distintas predicciones son los modelos de regresión, con los que podemos valorar prácticamente cualquier activo. Los modelos de regresión lineal múltiple nos permiten predecir el valor de un activo¹⁸ a través de una serie de variables relacionadas con dicho activo que consideramos tienen influencia en su precio. La formulación general de un modelo de regresión lineal múltiple se muestra a continuación.

$$Y = \beta_0 + \sum_1^i \beta_n * X_n + \mu \quad [1]$$

Donde:

Y -> Variable independiente (Valor del activo)

¹⁷ Marcías Guillén (2015)

¹⁸ Morales & Martínez (2021)

X -> Variable dependiente (Variables influyentes en el precio)

β -> Parámetros (Peso de la variable dependiente en la predicción del valor del activo)

ε -> Error existente

En ocasiones las relaciones existentes entre la variable independiente y las variables dependientes no tiene un desarrollo lineal, lo que implica efectos marginales¹⁹ no constantes. En dichos casos, aparecen transformaciones en las variables dependientes para representar mejor la relación existente. Las transformaciones más recurrentes son variables logarítmicas, variables elevadas al cuadrado o interacciones entre variables. Estas transformaciones nos permiten evitar uno de los mayores inconvenientes de esta técnica, su falta de eficacia al encontrarnos ante variables que no se relacionan linealmente con el precio del activo.

Existen multitud de estudios y proyectos en los que se han aplicado modelos de regresión para la valoración de activos obteniendo valoraciones precisas y con un coste inferior a otros modelos como las redes neuronales mencionadas anteriormente. A continuación se destacan las conclusiones más relevantes de los modelos de regresión analizados:

- Las variables independientes seleccionadas tienen una importancia²⁰ vital en la eficacia del modelo. Por ello, es necesario elaborar una lista con aquellas variables consideradas relevantes y analizar cuales se seleccionan finalmente para minimizar el error del modelo.
- La rigidez del modelo juega un papel importante a la hora de extrapolar el modelo a otras bases de datos. El objetivo es desarrollar un modelo que sea

¹⁹ Martínez de Ibarreta et al. (2019)

²⁰ En los modelos propuestos por Bellver & Martínez (2007) y Blasco (1998) podemos comprobar como la selección de las variables independientes desarrolla un papel vital en parámetros de interés del modelo que miden su precisión y error.

capaz de predecir correctamente el valor del activo sin sobreajustarlo en exceso a la base de datos que utilizamos para su creación, lo que generaría un problema²¹ conocido como “overfitting”, obteniendo un modelo que no da buenos resultados al extrapolarlo. Por el contrario, el modelo tampoco debe ser demasiado flexible²² de manera que considere como ruido parte de los datos y las predicciones no tengan la precisión esperada.

- Los modelos basados en una regresión tienen una mayor eficacia cuanto más centrados están en un área²³, es decir, un modelo²⁴ elaborado para predecir el valor de los bienes inmuebles residenciales de Nigeria ofrece mayor precisión y menor error que un modelo especificado para predecir el valor de bienes inmuebles en todo el territorio africano. Por ello, en multitud de ocasiones es aconsejable separar en base a unos parámetros la industria o sector que se pretende analizar y desarrollar un modelo para cada industria o sector en específico, con el objetivo de aumentar la precisión y reducir el error de los modelos.

Centrándonos en la industria del arte, hay que destacar que en este sector la existencia de variables cuantificables y observables²⁵ es, a diferencia de otras industrias, muy baja, por lo que debe recurrirse a criterios estéticos y variables más subjetivas²⁶ para obtener modelos de valoración relevantes. Esto supone un reto a la hora de introducir este tipo

²¹ Koehrsen (2018)

²² Concepto conocido como “underfitting.”

²³ En el estudio publicado por Guillén (2015) se pueden observar como la creación de un modelo centrado únicamente en obras del autor Pablo Picasso genera una mayor precisión en las predicciones y reduce el error de las mismas. Este mismo resultado se puede comprobar también en el estudio publicado por Núñez, Rey & Caridad y Ocerin (2017) al centrar el modelo de predicción de bienes inmuebles en la ciudad de Córdoba.

²⁴ Boluwatife & Chan (2018)

²⁵ Aznar & Guijarro (2007)

²⁶ Paleta de colores, nitidez, estilo, claridad u oscuridad.

de variables en el modelo y en muchas ocasiones resulta interesante²⁷ introducir dichas variables en forma de intervalo de valores en lugar de valores fijos.

Continuando con el análisis de la industria del arte, cabe destacar que en multitud de ocasiones las casas de subastas muestran sesgos sistemáticos en los precios de salida de las obras artísticas, es decir, muestran patrones sistemáticos que sobrevaloran o infravaloran el precio de salida de las obras, lo cual quedó evidenciado tras el desarrollo un modelo²⁸ que podía predecir qué obras iban a mostrar este sesgo antes incluso de conocer el precio de salida en las casas de subastas. Todo ello, nos muestra como no se está aprovechando la potencia de los avances tecnológicos y los modelos predictivos en esta industria, lo que permitiría a los intermediarios poder apoyarse en estas técnicas y mejorar sus estimaciones, de la misma forma que reducir coste y esfuerzo al disponer de una estimación desde la que poder basarse para establecer el precio de salida.

Siguiendo esta línea de aplicación de la tecnología a la industria del arte, un estudio²⁹ concluyó que aplicar el análisis mediante “machine learning” a la industria del arte puede traer enormes ventajas para todos sus participantes ya que se pudieron identificar una serie de tendencias o condiciones bajo las cuales los precios de remate aumentaban. Las tendencias más destacables fueron el mayor precio de aquellas obras con blanco, gris o negro como colores dominantes, obras con figuras más redondeadas y que no muestran bordes afilados o la tendencia a que las obras más caras se vendan en el mismo espacio temporal en el que se producen exhibiciones artísticas exitosas. Adicionalmente, también se resaltó las extensas oportunidades y campos de

²⁷ Bellver & Martínez (2007)

²⁸ Aubry et al. (2020)

²⁹ Hosny (2017). Un proyecto conocido como “The Green Canvas” evaluó si en una industria tan marcada por su carácter subjetivo, la tecnología podía ayudar a identificar tendencias o patrones que relacionasen el precio de venta con diversos atributos de las obras artísticas, con el objetivo de identificar aquellas variables que incrementasen los precios de remate de dichas obras.

investigación existentes para la aplicación de estas tecnologías en la industria del arte frente a otras más exploradas como los mercados bursátiles o los bienes inmuebles.

Finalmente, destacar un modelo³⁰ enfocado a la predicción del valor de obras de arte basado en una regresión hedónica mejorada añadiendo información de aquellas obras que ya habían sido vendidas en el pasado, junto con una técnica conocida como “*weighted model averaging*”³¹. En dicha técnica se emplean varios modelos con distintas variables explicativas y finalmente, se toma la media de los resultados de los modelos como la estimación final. Esta técnica nos permite disponer de la información de multitud de variables y no tener que hacer una selección de las más representativas. Se pudo comprobar como la adición de estas dos técnicas al modelo de regresión hedónica mejoró sustancialmente los resultados en cuanto a precisión y grado de explicación del modelo se refiere.

Por otro lado, todas las estimaciones y conclusiones extraídas de los modelos están condicionadas a la calidad en la especificación de estos, junto con los datos empleados para su creación. Por todo ello, también se han analizado libros y estudios que muestran los defectos más comunes a la hora de la formulación del modelo y las condiciones que deben cumplirse. Los modelos elaborados deben cumplir ciertas hipótesis básicas³² sobre las perturbaciones aleatorias, las variables explicativas y las especificaciones del modelo que nos permitan asegurar que existe una validación estadística, es decir, el modelo es significativo en su conjunto y las variables explicativas son significativas individualmente.

³⁰ Hodgson & Galbraith (2018)

³¹ Esta técnica permite que se evite tener que seleccionar un único modelo.

³² Martínez de Ibarreta et al. (2019)

Por último, otra técnica³³ muy dominante en mercados en los que no se dispone de una gran cantidad de datos para elaborar modelos predictivos, consiste³⁴ en apoyarse en precios de venta recientes de activos similares al que se pretende valorar, junto con algún ajuste por la diferencia temporal o por diferencias en los atributos o características de dicho activo. Esta técnica ha sido empleada con frecuencia en la industria del arte.

Siguiendo esta línea, ciertas industrias en las que la valoración de activos depende de multitud de suposiciones y variables subjetivas, se acude a expertos para que estimen estos parámetros y nos ofrezcan valoraciones que otras técnicas no son capaces de generar. Esta técnica ha sido empleada en la industria del arte³⁵ durante muchos años por las características que la industria presenta con multitud de variables de carácter subjetivo y la escasez de investigación en la valoración de obras de arte frente a otras industrias como las mencionadas anteriormente. No obstante, el desarrollo que ha experimentado el arte³⁶ en cuanto al cambio de perspectiva hacia un activo enfocado en la inversión y el incremento en la disponibilidad de información, han llevado a que estén surgiendo numerosos modelos más matemáticos y basados en las técnicas mencionadas anteriormente.

Conclusiones

En el presente capítulo se han analizado las principales técnicas de predicción del valor de activos. En un primer lugar, nos centramos en mercados en los que se dispone de gran cantidad de información, debido a que la entrada de la tecnología tuvo lugar

³³ Hodgson & Galbraith (2018)

³⁴ Esta técnica suele combinarse con otras de las mencionadas anteriormente.

³⁵ La mayoría de las casas de subastas disponen de un servicio en el que un experto tasador informa del valor aproximado de la obra en cuestión.

³⁶ Los estudios de predicción del valor de obras artísticas datan de fechas más recientes, lo que nos indica que la aplicación en la industria del arte de las diversas técnicas de predicción disponibles es mucho más reciente y menos extensa en comparación con otras industrias como los bienes inmuebles o los activos financieros.

numerosos años atrás, con el objetivo de entender cada técnica en profundidad. Adicionalmente, analizamos la información al respecto de la aplicación de cada técnica de predicción en la industria del arte.

Las redes neuronales aportan numerosos beneficios³⁷ en cuanto a precisión se refiere. Los modelos basados en esta técnica representan mejor³⁸ las relaciones no lineales generando un error menor en las predicciones. No obstante, ofrecen una serie de desventajas³⁹ que se acentúan en la industria del arte debido a la opacidad que ofrecen las redes neuronales y que otras técnicas como las regresiones hedónicas o la valoración por comparación nos permiten evitar.

Respecto a la industria del arte, la aplicación de modelos predictivos genera numerosas ventajas para todos los participantes en dicha industria ya que permite obtener un beneficio a los vendedores mediante la reducción costes en la tasación de las obras, junto con la reducción de esfuerzo y la mejora en la precisión⁴⁰ de los expertos tasadores. No obstante, el mayor inconveniente de la aplicación de modelos matemáticos a la industria del arte reside en la subjetividad característica de las variables que definen esta industria, lo que dificulta la forma en la que dichas variables subjetivas son introducidas en el modelo como datos numéricos.

³⁷ Boluwatife & Chan (2018)

³⁸ Núñez, Rey & Caridad y Ocerin (2017)

³⁹ Rodríguez (2021). Un tiempo de aprendizaje elevado, la necesidad de una elevada cantidad de datos y disponer de un operario que interprete la salida proporcionada por el modelo.

⁴⁰El estudio publicado por Aubry et al. (2020) nos muestra la existencia de sesgos sistemáticos en la tasación de obras artísticas.

En el siguiente capítulo se procede con el diseño del modelo predictivo que comenzará con la selección del tipo de obras artísticas objeto de estudio⁴¹, la selección de las principales técnicas en las que se basará nuestro modelo, junto con la identificación de las variables independientes que el proyecto busca analizar y que consideramos influyentes en el precio de las obras artísticas de autor.

3. DISEÑO DE UN MODELO DE PREDICCIÓN DE OBRAS DE ARTE

3.1. Justificación de la selección del área de desarrollo del modelo predictivo

Como se ha podido comprobar en el estado del arte, la industria del arte cuenta con un menor número de estudios e investigaciones que enfoquen este mercado desde la perspectiva de una de las nuevas clases de inversión alternativa. Esto ha provocado mayores dificultades en la identificación de las técnicas que ofrecen mejores resultados, las variables más determinantes en el precio de las obras o si realmente es eficaz aplicar estas técnicas en una industria como la artística que a priori parece completamente contraria al mundo de los modelos matemáticos. Por ello, uno de los objetivos del proyecto consiste en continuar aportando información e investigando en la aplicación de modelos predictivos a la industria del arte

Adicionalmente, como se observó en el estado del arte, las casas de subastas muestran sesgos sistemáticos⁴² en los precios de salida de las obras artísticas, es decir, muestran patrones sistemáticos que sobrevaloran o infravaloran el precio de las obras. Por ello, sería de gran utilidad poder avanzar en la investigación de modelos predictivos que permitan ayudar a todos los participantes en el mercado del arte, eliminando estos sesgos y pudiendo realizar valoraciones de mayor precisión, con un menor coste y esfuerzo, evitando en multitud de ocasiones tener que acudir a expertos tasadores.

⁴² Aubry et al. (2020)

Por otro lado, hemos visto como el nombre del autor juega un papel decisivo a la hora predecir el valor de una obra artística. Habitualmente podemos observar como las obras o cuadros de artistas de renombre alcanzan altos precios en las subastas con independencia de las características de dicha obra artística y sin seguir un patrón o método de valoración. Todos los estudios e investigaciones analizados de predicción del valor de obras artísticas se centran en autores de gran renombre como Pablo Picasso o Claude Monet. En los estudios analizados en el estado del arte, se ha comprobado como las obras de estos artistas muchas veces no siguen patrones lógicos y alcanzan precios extremos. La predicción del valor de estas obras está altamente correlacionada con la fama de su autor y la existencia de coleccionistas que quieran pagar grandes sumas de dinero por obras de estos autores. En numerosas ocasiones, dichas obras alcanzan precios desorbitados por cualidades muy particulares como ser una de las últimas obras en venta del artista o estar relacionadas con temas de la actualidad. En plena pandemia del COVID-19, el aclamado artista Banksy ha vendido una obra en la que se ve un niño jugando con una enfermera ‘superheroína’ por cerca de 19 millones de euros, lo que muestra la irracionalidad del precio⁴³ que alcanzan este tipo de obras.

No obstante, no se ha encontrado ningún estudio que se enfoque en las obras de autor, aquellas obras cuyos artistas no son tan conocidos. Este tipo de obras componen el grueso del mercado del arte, ya que aunque el público menos introducido en este mundo solo escuche hablar de los momentos puntuales en los que se venden obras de autores famosos por un elevado precio, el día a día de las casas de subastas y del mercado del arte gira entorno a pinturas de artistas con menor renombre. Uno de los objetivos de este proyecto es determinar si en este tipo de obras el precio queda fijado en mayor medida por los atributos y por la calidad de la obra en cuestión, o si se mantiene la volatilidad que caracteriza los estudios analizados en el estado del arte de obras de autores famosos.

⁴³ En esta ocasión, vemos como esta obra de Banksy ha alcanzado un precio tan elevado debido a la temática empleada en ella lo que ha llamado la atención de numerosos coleccionistas.

Por todo lo mencionado anteriormente, en el presente proyecto desarrollaremos un modelo predictivo de inversión en obras de arte de autores no tan conocidos con el objetivo de poder seguir investigando y aportando información en este campo sobre la utilidad de la aplicación de técnicas predictivas en la industria del arte, analizar la posibilidad de facilitar el trabajo a los expertos tasadores mediante la reducción de costes y esfuerzo, junto con la mejora de la precisión en las predicciones.

Adicionalmente, desarrollamos un modelo centrado en un tipo de obras que no han sido analizadas con el objetivo⁴⁴ de desarrollar un modelo que sirva de base para que los expertos y las casas de subastas puedan partir de una estimación inicial del precio y posteriormente realizar los ajustes convenientes⁴⁵ que el modelo no pueda captar.

3.2. Selección de las técnicas del modelo predictivo

En la actualidad existen diversas técnicas a la hora de predecir el valor de un activo, las cuales han sido explicadas en el estudio del arte. No obstante, realizaremos un rápido resumen de aquellas que podrían ser aplicadas en la valoración de obras artísticas y analizaremos cada técnica con el fin de seleccionar el conjunto de técnicas que se ajusten mejor a nuestro modelo predictivo de inversión en obras de arte justificando las razones o inconvenientes de la selección de dichas técnicas.

⁴⁴ Bajo la suposición de que encontremos un modelo con una buena capacidad explicativa que facilite la labor de los tasadores. La inexistencia de estudios que analicen esta área de la industria del arte no nos permite conocer la eficacia de aplicar dichas técnicas predictivas en ella.

⁴⁵ Ajustes debido a tendencias actuales de ciertos estilos o colores que estén más demandados en la actualidad.

Por un lado, existen técnicas más avanzadas tecnológicamente que están aumentando su popularidad y aplicación en muchas áreas en las que se dispone de gran cantidad de datos junto con variables objetivas y cuantificables, entre ellas destaca las redes neuronales como la más potente y popular. Pese a las numerosas ventajas que ofrece esta técnica⁴⁶, consideramos que su aplicación no es apropiada para la industria del arte por las razones que explicamos a continuación:

- Las redes neuronales son muchas veces calificadas como “black box”, esto se debe a que se trata de una tecnología la cual carece de una explicación⁴⁷ visible detrás de las valoraciones que ofrece el modelo. Una vez la red neuronal ha sido entrenada y comienza su utilización para la predicción, no se conoce el funcionamiento interno del modelo ni el criterio para la obtención de los resultados. Por las características de la industria del arte, consideramos que es necesario un modelo transparente en el que se puedan identificar las variables explicativas así como sus pesos en la predicción final. De esta forma, sirve como herramienta para que alguien introducido en la industria pueda usar dicho modelo y combinarlo con su conocimiento de cara a realizar ajustes en la predicción final o en el modelo para representar cambios en las tendencias del mercado o matices que el modelo puede no tener en cuenta dada su generalidad, lo que haría aún más efectiva la utilización del modelo.
- Siguiendo la línea de lo mencionado en el punto anterior, la industria del arte se encuentra en constante evolución, surgiendo nuevas tendencias o cambio en los gustos que modifican drásticamente el peso de ciertas variables y finalmente, el valor de las obras artísticas. Por ello, creemos que es de vital importancia un modelo transparente como el mencionado anteriormente. Creemos que esta combinación de la predicción del modelo y el conocimiento de un experto se

⁴⁶ Mayor precisión, menor error, mejor adaptación a relaciones no lineales. Véase estado del arte para más información.

⁴⁷ Burke & Burns (2021)

adapta mucho mejor a la industria del arte y proporciona mejores resultados frente a un modelo opaco basado en redes neuronales.

- Adicionalmente, las redes neuronales trabajan con información numérica, por lo tanto todos los problemas deben ser traducidos a valores numéricos⁴⁸ antes de introducirlos en la red neuronal, lo que supone un problema para nuestro modelo ya que la existencia de variables explicativas de carácter subjetivo dificulta en gran medida su traducción a valores numéricos fijos. Como se vio en el estado del arte, las ventajas de la introducción de este tipo de variables en intervalos⁴⁹ en vez de valores fijos.

A continuación se presenta un diagrama de flujos del funcionamiento de un modelo predictivo general basado en redes neuronales. En él se puede comprobar la falta de transparencia que caracteriza a esta técnica y que no es adecuada para la industria del arte. La capa oculta puede estar conformada por varias capas.

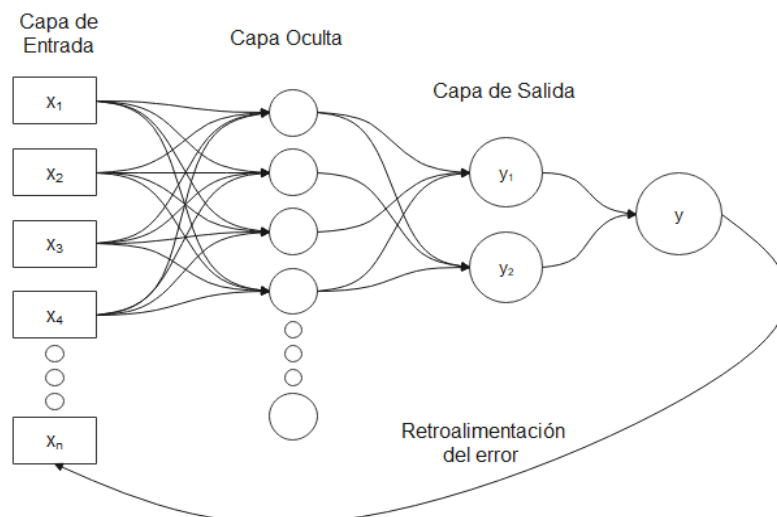


Ilustración 3. Fuente: Elaboración Propia (2021)

⁴⁸ Burke & Burns (2021)

⁴⁹ Bellver & Martínez (2007)

Por otro lado, los modelos de regresión hedónica o modelos hedónicos son ampliamente utilizados en la valoración de todo tipo de activos, ofreciendo resultados relativamente precisos a un bajo coste. Esta técnica es muy popular y posee mucha solidez ya que lleva muchos años utilizándose en multitud de áreas y son innumerables los estudios de investigación que se han realizado a su alrededor, comprobando su validez y optimizando su eficacia. En el estado del arte se ha comprobado cómo es una técnica que ofrece una alta precisión⁵⁰ y que ofrece otras ventajas frente a las redes neuronales como un menor coste, tanto económico como temporal, ya que no se requiere una base de datos tan grande como para un modelo basado en redes neuronales y también se puede entrenar el modelo en un menor tiempo y con un menor coste computacional.

En el siguiente diagrama de flujos se representan las fases de trabajo para la correcta elaboración de un modelo de predicción basado en una regresión.

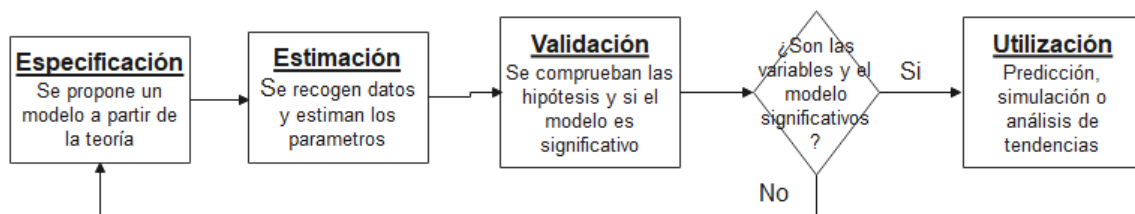


Ilustración 4. Fuente: Elaboración Propia (2021)

Los modelos basados en una regresión hedónica son muy utilizados a la hora de predecir el valor de bienes inmuebles. La industria del arte comparte multitud de características con el mercado de los bienes inmuebles, ya que en ambas existe un mercado para activos que se comercian de manera poco frecuente y dichos activos son únicos pero a la vez mantienen una serie de atributos de manera común. Ello nos indica que puede ser una técnica efectiva para la predicción de obras artísticas.

⁵⁰ Hodgson & Galbraith (2018)

A continuación, se presenta la formulación y el diagrama de flujos de un modelo general basado en una regresión hedónica. Posteriormente se realizará un diagrama de flujos más detallado de nuestro modelo en concreto.

$$Y = X * B + E \quad [2]$$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ 1 & x_{31} & \dots & x_{3k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & x_{nk} \end{pmatrix} * \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad [2]$$

Donde:

- **Y → Vector de variables dependientes o endógenas.** En los modelos de predicción es el precio del activo que buscamos determinar.
- **X → Matriz de variables independientes o exógenas.** En los modelos de predicción se trata del conjunto de variables explicativas del precio del activo bajo estudio.
- **B → Vector de parámetros de la regresión múltiple.** Los parámetros son indicadores de la importancia que las variables independientes a las que acompañan tienen en el precio del activo bajo estudio. Se determinan a través de diferentes métodos que se verán en capítulos posteriores.
- **E → Vector de perturbaciones aleatorias de la regresión múltiple.** Reflejan todos los cambios que experimentan las variables dependientes y no quedan reflejados mediante los cambios en las variables dependientes.

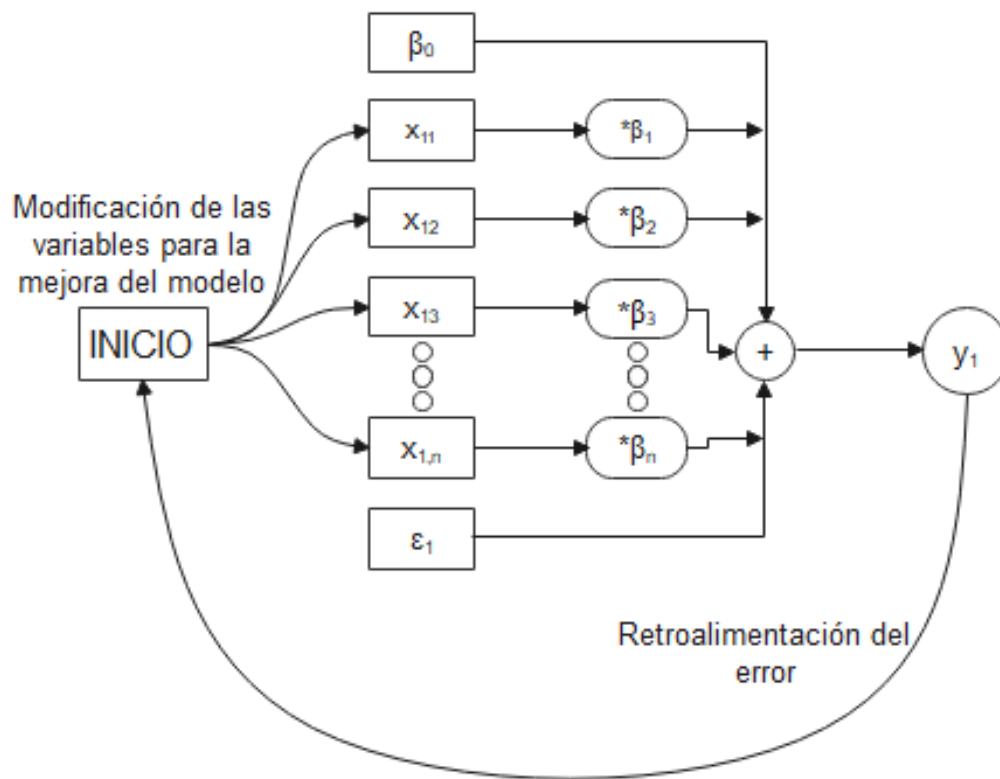


Ilustración 5. Fuente: Elaboración Propia (2021)

Si nos centramos en la industria del arte, consideramos que esta técnica se adapta perfectamente a las características de dicho sector ya que es una técnica completamente transparente y se pueden conocer y seleccionar las variables explicativas, los pesos de dichas variables en la valoración final y ofrece la posibilidad de realizar ajustes⁵¹ en el modelo cuando sea necesario.

Por todo lo mencionado, esta será la técnica principal en la que se basará nuestro modelo predictivo.

⁵¹ Cambios en las variables explicativas o en los pesos de las mismas.

Adicionalmente, cuando existe información de activos que han sido vendidos en repetidas ocasiones en el pasado, podemos basar las estimaciones en dichos precios de venta del pasado y ajustarlos en función de los cambios que el mercado haya experimentado desde dichas ventas. A continuación se muestra la formulación y el diagrama de flujos de un modelo basado en la técnica de ventas pasadas.

$$\mathbf{P} = \mathbf{X} * \mathbf{B} + \mathbf{E} \quad [3]$$

$$\begin{pmatrix} p_{1,t} \\ p_{2,t} \\ p_{3,t} \\ \vdots \\ p_{n,t} \end{pmatrix} = \begin{pmatrix} 1 & p_{(t-q),1} & \dots & x_{1,k} \\ 1 & p_{(t-q),2} & \dots & x_{2,k} \\ 1 & p_{(t-q),3} & \dots & x_{3,k} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & p_{(t-q),k} & x_{n,1} & x_{n,k} \end{pmatrix} * \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad [3]$$

Donde:

- **P → Vector de variables dependientes o endógenas.** Precio de venta del activo que buscamos determinar.
- **X → Matriz de variables independientes o exógenas.** Conjunto de variables explicativas del precio del activo bajo estudio. La primera variable explicativa es $p_{(t-q),i}$, la cual representa el precio de venta del activo en el pasado⁵².
- **B → Vector de parámetros de la regresión múltiple.** Los parámetros son indicadores de la importancia que las variables independientes a las que acompañan tienen en el precio del activo bajo estudio. Se determinan a través de diferentes métodos que se verán en capítulos posteriores.
- **E → Vector de perturbaciones aleatorias de la regresión múltiple.** Reflejan todos los cambios que experimentan las variables dependientes

⁵² q meses desde el momento t.

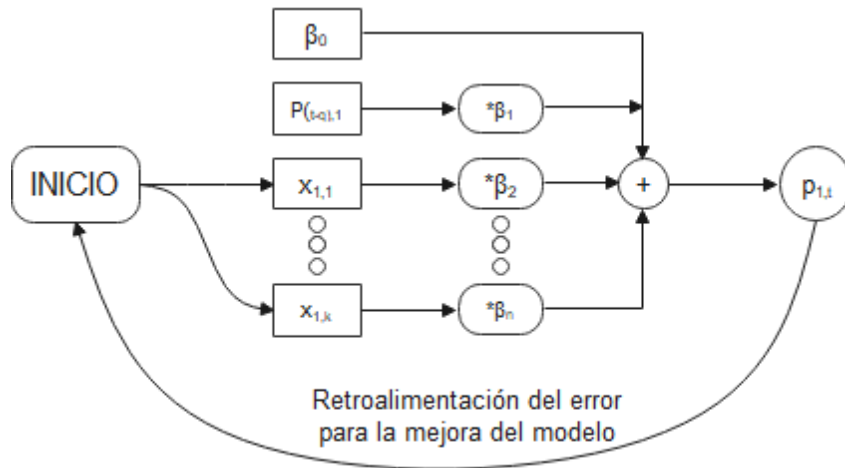


Ilustración 6. Fuente: Elaboración Propia (2021)

Este método es aplicado en numerosas ocasiones por las casas de subastas a la hora de determinar el precio de salida de las obras artísticas. Es una técnica predictiva que ofrece resultados precisos⁵³ en el mercado del arte, no obstante, ofrece el inconveniente de ser útil únicamente en aquellos casos en los que la obra artística haya sido vendida en el pasado y se conozca dicho precio venta y su fecha correspondiente. Por ello, la introduciremos en nuestro modelo siempre que cumpla la condición mencionada y dispongamos de dicha información.

Finalmente, el método de valoración más común y el que más años lleva en uso en la industria del arte consiste en acudir a la valoración de expertos. Las casas de subastas disponen de tasadores expertos. Cualquier individuo puede acudir con una obra artística y dichos tasadores realizan una estimación del precio al que será vendida la obra. No obstante, para obras artísticas que no se esperan vender por un precio elevado ya que no

⁵³ Generalmente cuanto más reciente es la fecha de venta de la obra, más precisa es la estimación obtenida con esta información.

proviene de un autor famoso, dicho servicio puede ser sustituido por un modelo predictivo que permita ahorrar costes y tiempo tanto al propietario de la obra como a la casa de subastas o al tasador. Por ello, consideramos que la mejor opción de predicción del valor de obras artísticas de autor consiste en la combinación entre un modelo predictivo y el conocimiento de un experto. De esta manera la casa de subastas dispone del modelo que puede ser actualizado por los expertos en función de las tendencias actuales del mercado.

Los estudios realizados hasta la fecha muestran que las estimaciones más precisas se consiguen cuando se dispone de un modelo⁵⁴ para cada uno de los autores. No obstante, en este proyecto tratamos de elaborar un modelo de predicción para obras de autor en las que el artista raramente tiene influencia en el precio de la obra. No se han encontrado estudios que analicen este sector de la industria por lo que consideramos que lo adecuado es plantear un modelo general ya que en multitud de ocasiones se analizan obras artísticas que no están firmadas, por lo que no se conoce el autor, o sus autores no son conocidos ni influyentes en el precio de la misma.

En resumen, elaboraremos un modelo predictivo basado en una regresión hedónica añadiendo la información de las ventas pasadas cuando se disponga de ellas. Todo ello con el objetivo de que el modelo sea empleado por las casas de subastas de manera que los expertos tasadores se basen en sus estimaciones y realicen los cambios oportunos a las mismas para reflejar aquello que el modelo no permita captar y puedan mantener el modelo actualizado en función de las tendencias del mercado.

⁵⁴ Se vio en el estado del arte al analizar el estudio de predicción de obras artísticas de Pablo Picasso (Marcías Guillén (2015)). Este mismo resultado también se observó en otros estudios como al analizar el precio de venta de bienes residenciales en la ciudad de Córdoba (Núñez, Rey & Caridad y Ocerin (2017))

En los siguientes capítulos procederemos con la selección de las variables objeto de estudio y la especificación del modelo en el que se basará este estudio.

3.3. Selección del conjunto de variables explicativas

Hasta la actualidad, los estudios existentes que han tratado de crear modelos de predicción para obras de arte comparten una serie de variables que están presentes en la gran mayoría de los estudios analizados. Entre las variables más influyentes a la hora de calcular una estimación del precio de la obra en cuestión y cuya aplicación tenga sentido en nuestro modelo, destacamos las siguientes:

- **Precio de venta.** Variable dependiente y cuantitativa continua medida en euros. Precio final al que se vendió la obra artística por la casa de subastas. Se trata de la variable que buscamos explicar a partir de las variables que se mencionan a continuación. El modelo elaborado en este estudio se centra en aquellas obras cuyo precio de venta no sobrepasa los 6.000€ – 7.000€. El precio de venta de la obra j queda codificado como ω_j en el modelo de predicción.
- **Tamaño de la obra.** Variable cuantitativa continua medida en pulgadas al cuadrado. Analizaremos esta variable mediante el área de la obra. Dentro de esta variable, debemos tener en cuenta que, generalmente, las obras abstractas suelen tener mayor tamaño que las obras figurativas. El tamaño de la obra j queda codificado como T_j en el modelo de predicción.
- **Siglo de creación.** Variable cuantitativa discreta medida por el siglo en el que fue pintada la obra ya que en muchas ocasiones resulta complicado determinar el año de creación. Las obras se revalorizan con el paso del tiempo, por lo que las obras de mayor antigüedad tienden a mostrar un mayor precio que las obras que

han sido creadas recientemente. El siglo de creación de la obra j queda codificado como S_j en el modelo de predicción.

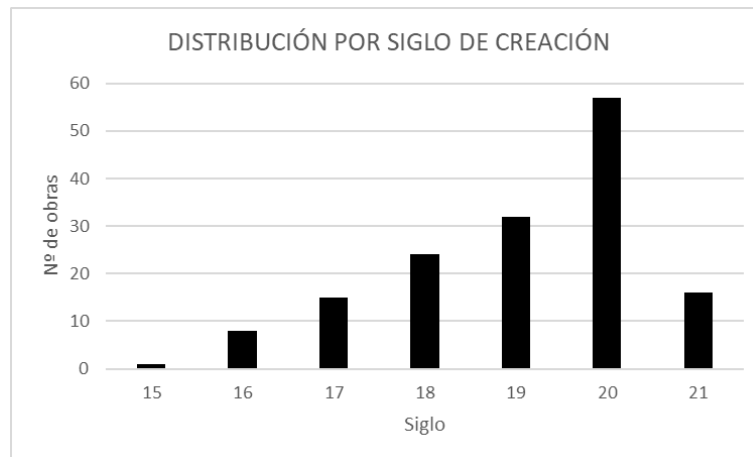


Ilustración 7. Fuente: Elaboración Propia (2021)

- **Origen⁵⁵**. Variable nominal. Dentro de esta variable se identifican cuatro categorías: obras de origen americano, europeo, asiático y africano. A la hora de formular el modelo introduciremos cada origen como una variable independiente de manera que podamos conocer a través del parámetro que acompaña a cada variable, la influencia que cada origen tiene en el precio de venta. El origen de la obra j queda codificado como $P_{i,j}$ en el modelo de predicción. Se trata de una variable nominal dicotómica en la que i representa un origen en concreto.

⁵⁵ Valores que toma la variable: América (1), Europa (2), Asia (3) y África (4).



Ilustración 8. Fuente: Elaboración Propia (2021)

- **Casa de subastas**⁵⁶. Variable cualitativa nominal. Se han seleccionado cinco casas de subastas que tienen un gran volumen de ventas: Clars Auction Gallery, Jasper52, ArtForSale.com, Royal Treasures Warehouse LLC y Hudson Valley Auctions. Al igual que la variable origen, cada casa de subastas se introduce en el modelo como una variable independiente con el mismo fin explicado en la anterior variable. La casa de subastas en la que se vendió la obra j queda codificado como $C_{x,j}$ en el modelo de predicción. Se trata de una variable nominal dicotómica en la que x representa una casa de subastas en concreto.

⁵⁶ Valores que toma la variable: Clars Auction Gallery (1), Jasper52 (2), ArtForSale.com (3), Royal Treasures Warehouse LLC (4) y Hudson Valley Auctions (5).



Ilustración 9. Fuente: Elaboración Propia (2021)

A continuación, presentamos diversas variables de carácter más subjetivo que consideramos pueden tener influencia en la valoración de obras de autor. Como comentamos anteriormente, dada la inexistencia de estudios que se centrasen en el mercado de obras artísticas de autor, estas variables son experimentales, tratando de analizar cuáles son representativas para obtener un modelo preciso y con una buena capacidad explicativa. Como se ha mencionado a lo largo del proyecto, este estudio busca valorar las obras por sus características visuales, por lo que las variables que se describen a continuación buscan enmarcar este tipo de características de las obras artísticas. Posteriormente, estudiaremos cuan significativas son estas variables.

- **Paleta de colores**⁵⁷. Variable cualitativa dicotómica. Diferenciamos las obras artísticas entre aquellas que presentan una gran variedad de colores y aquellas realizadas en un color determinado variando su tonalidad. La existencia de variedad o no de colores en la obra j queda codificado como O_j en el modelo de predicción. A continuación se muestra la diferencia mencionada:



Ilustración 10. Obra en un color.
Fuente: Elaboración Propia (2021)



Ilustración 11. Obra en varios colores. Fuente: Elaboración Propia (2021)

- **Nitidez**⁵⁸. Variable cualitativa dicotómica. Destacamos la diferencia entre aquellas obras que están perfectamente definidas y dibujadas con el máximo detalle y otras obras en las que se tiende a difuminar y mezclar los contornos de las figuras representadas. La nitidez o el difuminado de la obra j queda codificado como N_j en el modelo de predicción. A continuación se presentan dos obras en las que se pueden comprobar las diferencias mencionadas.

⁵⁷ La variable toma el valor 1 cuando la obra presenta una variedad de colores considerable y el valor 0 cuando predomina un color en diversas tonalidades.

⁵⁸ La variable tomo el valor 1 cuando la obra está representada de manera nítida y el valor 0 cuando la obra se representa de forma difuminada.



Ilustración 12. Obra con baja nitidez.
Fuente: Elaboración Propia (2021)



Ilustración 13. Obra con alta nitidez. Fuente: Elaboración Propia (2021)

- **Suavidad**⁵⁹. Variable cualitativa dicotómica. Dividimos las obras entre aquellas que muestran líneas rectas y vértices en punta y dan la sensación de ser más ásperas y por otro lado, otras obras con mayor número de líneas curvas, evitando los vértices y generando una apariencia más suave y redondeada. La suavidad de la obra j queda codificado como F_j en el modelo de predicción.



Ilustración 15. Obra suave. Fuente: Elaboración Propia (2021)



Ilustración 14. Obra no suave. Fuente: Elaboración Propia (2021)

⁵⁹ La variable toma el valor 1 cuando la obra muestra líneas curvas y suavidad y el valor 0 cuando predominan las líneas rectas y los vértices.

- **Material⁶⁰**. Variable cualitativa nominal. Dentro de esta variable distinguimos tres tipos de material empleados con frecuencia en la industria del arte: lienzo, papel y acuarela. Introduciremos cada tipo de material como una variable independiente del modelo. El material de creación de la obra j queda codificado como $M_{z,j}$ en el modelo de predicción. Se trata de una variable nominal dicotómica en la que z representa un material en concreto.
- **Estilo⁶¹**. Variable cualitativa nominal. Diferenciamos los siguientes tipos de estilos en las obras artísticas: retrato, paisaje, abstracto y surrealista; introduciendo cada estilo como una variable independiente del modelo. El estilo de la obra j queda codificado como $E_{w,j}$ en el modelo de predicción. Se trata de una variable nominal dicotómica en la que w representa un estilo en concreto.



Ilustración 17. Retrato/Escena.
Fuente: Elaboración Propia (2021)



Ilustración 16. Paisaje. Fuente: Elaboración Propia (2021)

⁶⁰ Valores que toma la variable: lienzo (1), papel (2) y acuarela (3).

⁶¹ Valores que toma la variable: paisaje (1), retrato (2), abstracto (3) y surrealista (4).



*Ilustración 18. Estilo surrealista.
Fuente: Elaboración Propia (2021)*



Ilustración 19. Estilo abstracto. Fuente: Elaboración Propia (2021)

- **Colores claros u oscuros⁶².** Otra variable cualitativa dicotómica cuya influencia en el precio de la obra nos gustaría analizar son la diferencia entre aquellas obras realizadas en colores claros que reflejan la luz, aportan luminosidad y crean sensación de amplitud frente a obras realizadas en colores oscuros que absorben la luz, creando sensación de estrechez y resaltando más. El predominio de colores claros u oscuros en la obra j queda codificado como D_j en el modelo de predicción.

⁶² La variable toma el valor 1 cuando la obra está representada en colores oscuros y el valor 0 cuando la obra se representa en colores claros.



Ilustración 21. Obra oscura. Fuente: Elaboración Propia (2021)



Ilustración 20. Obra clara. Fuente: Elaboración Propia (2021)

Finalmente, comentar la existencia de una serie de variables que son ampliamente empleadas en los modelos de predicción de obras artísticas analizados en el estado del arte. No obstante, consideramos que su implicación en nuestro modelo no es adecuada por las razones que se especifican a continuación:

- **Artista.** Dentro de los cuadros de autor, ciertos artistas tienen más fama que otros por lo que pueden generar variaciones en el precio. No obstante, nuestro modelo se ha limitado a obras que alcanzan un máximo de 6.000 € - 7.000 € con el objetivo de analizar aquellos cuadros en los que la importancia del autor es efímera y reducir al máximo la influencia de dicho autor.
- **Año del artista.** En muchas ocasiones la etapa de la carrera profesional en la que el artista realizó la obra tiene un papel fundamental en la determinación del precio de venta, por lo que esta obra aparece representada en multitud de estudios de predicción del valor de obras artísticas. Esta variable es fundamental en aquellos artistas de gran renombre. No obstante, la inclusión de dicha variable en nuestro modelo no tiene sentido ya que el autor o bien no es conocido por la falta de su firma o su influencia en el precio es inexistente.

3.4. Descripción estadística de la base de datos empleada para la elaboración del modelo predictivo

Para el desarrollo del modelo de predicción de obras de arte, se ha elaborado una base de datos⁶³ con las variables especificadas en el capítulo anterior. Dicha base de datos ha sido obtenida desde la plataforma online Live Auctioneers⁶⁴. A continuación realizamos una descripción y análisis estadístico de la base de datos e identificamos algunas relaciones de interés entre ciertas variables explicativas y el precio que surgen de un primer análisis superficial y buscan entender mejor las características de las obras artísticas a las que nos enfrentamos y realizar transformaciones en el modelo en el caso de que existan relaciones no lineales entre alguna de las variables del modelo y el precio de venta.

La base de datos consta de 153 observaciones (32 observaciones correspondientes a obras que han sido vendidas en multitud de ocasiones) con un precio medio de 1.122 € y una dimensión media de 676 pulgadas al cuadrado. Las observaciones corresponden a obras artísticas cuya creación oscila entre el siglo 15 y el siglo 21. Sus correspondientes fechas de venta oscilan entre 15/08/2004 y 30/05/2021.

Una de las hipótesis que nos planteamos en la selección de las variables es el incremento en el precio de venta conforme las obras son de mayor antigüedad. En el siguiente grafico se representa esta relación y la línea de tendencia entre ambas variables. Como se puede comprobar esta hipótesis se cumple por lo que el precio de venta y el siglo de creación de la obra artística mantienen una relación negativa. Esta relación se acentúa en obras cuya creación data entre el siglo 15 y el siglo 17.

⁶³ Véase anexo para más información de la base de datos.

⁶⁴ "Marketplace" de arte, antigüedades y artículos de colección más grande del mundo.

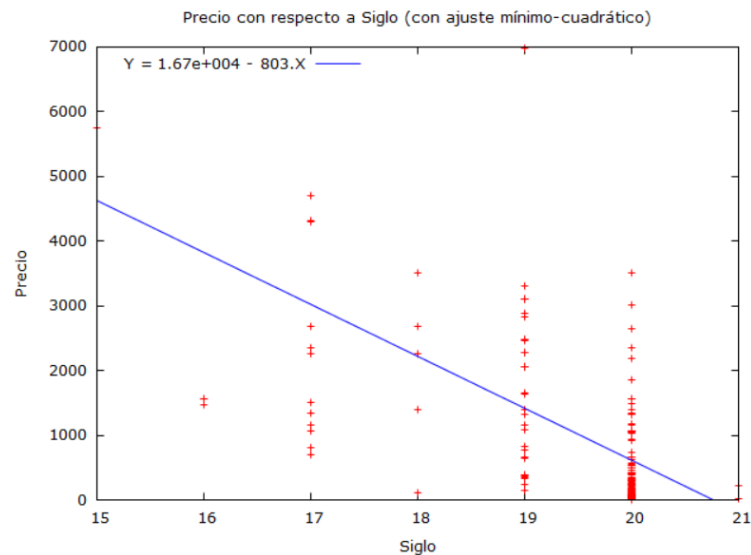


Ilustración 22. Fuente: Elaboración Propia (2021)

Otra de las relaciones que se analizan en esta base de datos es la que vincula las variables del precio de venta y el tamaño⁶⁵ de las obras artísticas. Pese a que se pueda pensar que las obras de mayor tamaño tienden a venderse a un precio mayor, esta relación no se mantiene en nuestra base de datos. El precio y el tamaño sostienen una relación neutra en la que no se aprecia una variación en el precio conforme aumenta o disminuye el tamaño.

⁶⁵ Medido a través del área en pulgadas al cuadrado.

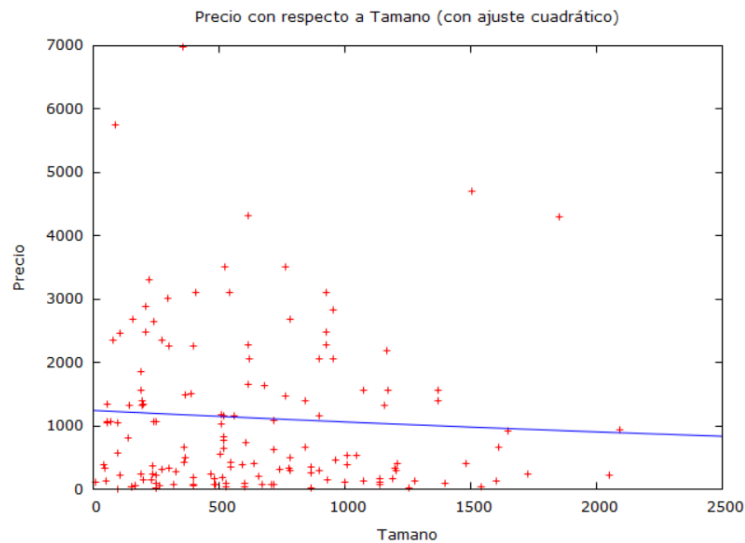


Ilustración 23. Fuente: Elaboración Propia (2021)

Adicionalmente, consideramos que las obras de estilo abstracto tienen un tamaño mayor que las obras de estilo figurativo. Esta hipótesis también se cumple al analizar nuestra base de datos. Las obras abstractas tienen un tamaño medio de 920 pulgadas al cuadrado frente a las 630 pulgadas al cuadrado de las obras figurativas. Por ello, en la especificación del modelo introduciremos una variable de interacción que se especificará en el siguiente capítulo.

3.5. Especificación del modelo de predicción

En este apartado se procede con la especificación inicial del modelo de predicción de obras de arte. En los apartados anteriores se han desarrollado las razones por las que el modelo se basará en una regresión hedónica junto con la información que nos aporten aquellas obras que han sido vendidas en varias ocasiones, así como las variables iniciales que se van a estudiar, sus relaciones con el precio de venta y un primer análisis estadístico de la base de datos empleada en el cálculo de los parámetros.

El modelo inicial de predicción de obras de arte se presenta a continuación. El precio de una obra artística j queda determinado de la siguiente forma:

$$\begin{aligned} \omega_{1,j} = & \beta_0 + \beta_1 * T_j + \beta_2 * S_j + \sum_{i=3}^{i=5} (\beta_i * P_{i,j}) + \sum_{x=6}^{x=9} (\beta_x * C_{x,j}) + \beta_{10} * O_j + \beta_{11} * N_j \\ & + \beta_{12} * F_j + \sum_{z=13}^{z=14} (\beta_z * M_{z,j}) + \sum_{w=15}^{w=17} (\beta_w * E_{w,j}) + \beta_{18} * D_j + \beta_{19} \\ & * (T_j * E_{16,j}) + \mu_j \end{aligned}$$

[4]

Donde:

$\omega_{1,j}$	Precio de venta (Regresión hedónica)	β_n	Parámetro de la regresión múltiple	$P_{i,j}$	Origen	$i=3 \rightarrow$ Europa $i=4 \rightarrow$ Asia $i=5 \rightarrow$ África
O_j	Paleta de colores	T_j	Tamaño (pulgadas ²)	$C_{x,j}$	Casa de subastas	$x=6 \rightarrow$ Jasper52 $x=7 \rightarrow$ ArtForSale.com $x=8 \rightarrow$ Royal Treasures Warehouse LLC $x=9 \rightarrow$ Hudson Valley Auctions
D_j	Claridad	S_j	Siglo	$E_{w,j}$	Estilo	$i=15 \rightarrow$ Paisaje $i=16 \rightarrow$ Abstracto $i=17 \rightarrow$ Surrealista
N_j	Nitidez	F_j	Suavidad	$M_{z,j}$	Material	$z=13 \rightarrow$ Papel $z=14 \rightarrow$ Acuarela

Tabla 1. Fuente: Elaboración Propia (2021)

Las variables cualitativas ordinales⁶⁶ de nuestro modelo se introducen como variables dicotómicas respecto a una categoría base, es decir, dichas variables toman el valor 1 en caso de que se cumpla esa condición determinada y 0 en todos los casos complementarios. Por ejemplo, la variable $M_{13,j}$ toma el valor 1 si la obra está realizada en papel y el valor 0 si la obra esta realiza sobre lienzo o en acuarela. Este tipo de variables las introducimos como variables dicotómicas respecto a la categoría base, cuya variable no introducimos. El significado de los parámetros consiste en la influencia que una determinada categoría respecto a la categoría base tiene en el precio. Por ejemplo, la variable $M_{13,j}$ va acompañada de un parámetro β_{13} que nos indica la influencia que tiene en el precio de venta de una obra que dicha obra sea creada en papel en lugar de en lienzo (categoría base en este caso). Las categorías bases en las variables cualitativas ordinales introducidas en nuestro modelo son las siguientes:

- Origen base $\rightarrow$ América
- Casa de subastas base $\rightarrow$ Clars Auction Gallery
- Material base $\rightarrow$ Lienzo
- Estilo base $\rightarrow$ Retrato/Escena

Se introduce una variable de interacción entre el tamaño y las obras de estilo abstracto. Esto se debe a que las obras de estilo abstracto suelen tener un mayor tamaño por lo que la influencia del tamaño en el precio de las obras de esta tipología difiere de las obras de otra tipología.

En aquellos casos en los que la obra haya sido vendida en el pasado y se conozca dicho precio de venta se acude al siguiente modelo:

⁶⁶ Origen, casa de subastas, material y estilo.

$$\omega_{3,j} = \alpha * \omega_{1,j} + (1 - \alpha) * \omega_{2,j} \quad [5]$$

$$\omega_{2,j} = \beta_0 + \beta_1 * \omega_{p,i} + \beta_2 * R_j + \mu_j \quad [6]$$

Donde:

$\omega_{1,j}$	Precio de venta (Regresión hedónica)	β_n	Parámetro de la regresión múltiple	R_j	N.º de Meses Transcurridos entre las ventas
$\omega_{2,j}$	Precio de venta (Regresión con venta pasada)	$\omega_{p,i}$	Precio de venta (Venta pasada)	α	Parámetro del modelo combinado
$\omega_{3,j}$	Precio de venta (Modelo combinado)				

Tabla 2. Fuente: Elaboración Propia (2021)

El valor óptimo de α se obtiene mediante observación. No obstante, la literatura nos dice que es difícil mejorar el peso igualitario de ambas observaciones, es decir, un valor para α igual a 0.5. Por esta razón, junto con la necesidad de tener que desarrollar una base de datos de forma manual debido a la dificultad de obtener una base de datos amplia con las características visuales que están siendo analizadas en este estudio, se otorga el valor 0.5 al parámetro α .

Finalmente, se presenta un diagrama de flujos con el modelo inicial de predicción de obras de arte de autor.

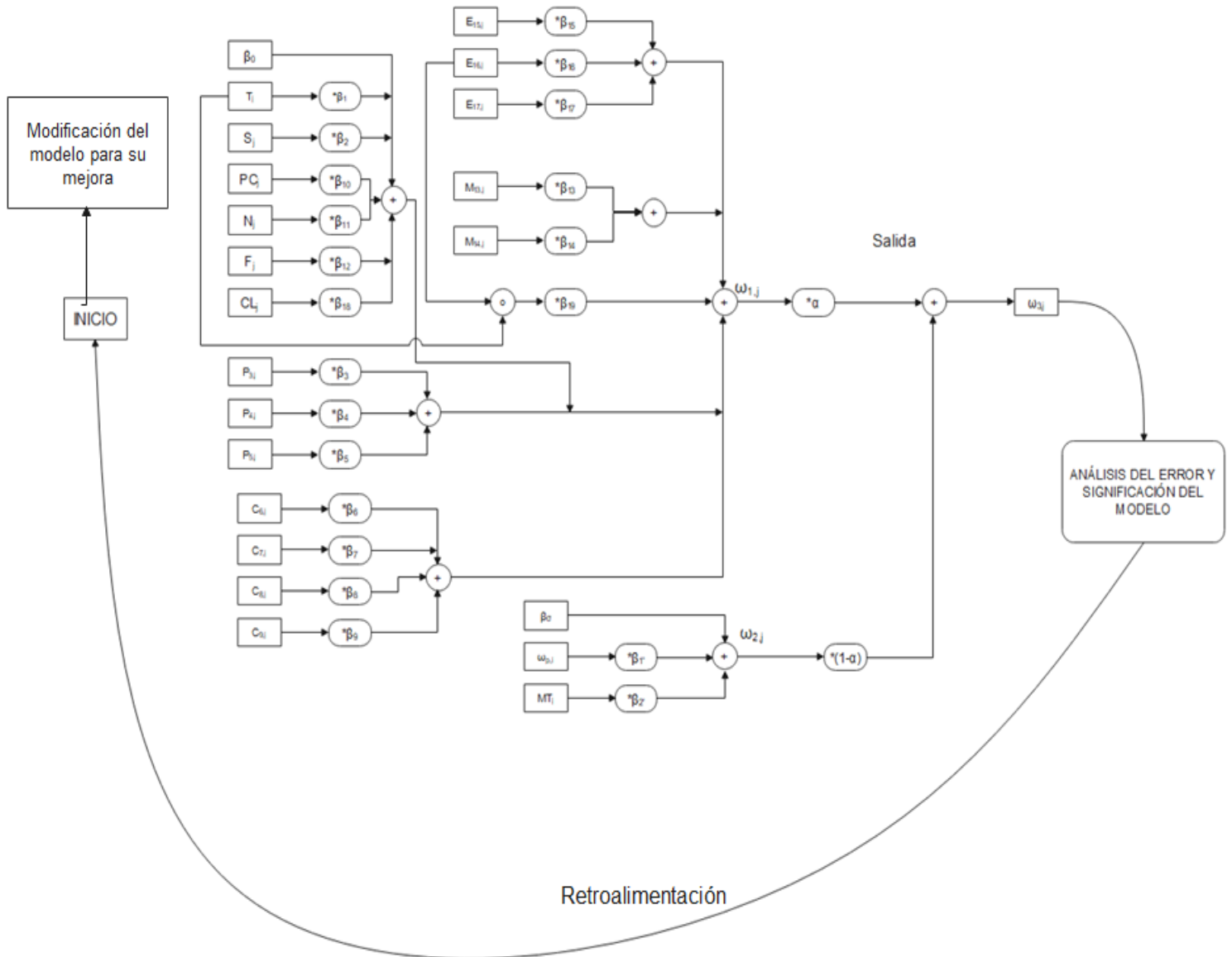


Ilustración 24. Fuente: Elaboración Propia (2021)

3.6. Conclusiones

En este capítulo hemos desarrollado el proceso para el diseño de un modelo inicial para la predicción de valor de obras de autor.

En un primer lugar, nos hemos decantado por centrar el modelo en artistas con poco renombre ya que supone el grueso y día a día de la industria del arte y no existe investigación que se centre en esta área de la industria.

Posteriormente, hemos analizados las técnicas de predicción disponibles en la actualidad evaluando su adaptación a nuestro modelo. La técnica seleccionada consiste en una regresión hedónica junto con la información de aquellas obras que han sido vendidas en multitud de ocasiones.

Mas adelante, hemos elaborado una lista con las variables independientes que consideramos son influyentes en el precio final de las obras.

Por último, se ha realizado una descripción y análisis de la base de datos empleada para el diseño del modelo y la especificación del modelo inicial para la predicción de obras de arte de artistas con menor renombre.

En el siguiente capítulo procederemos con el cálculo de los parámetros, el análisis del modelo y su validación mediante la aplicación GRETL.

4.VALIDACIÓN DEL MODELO DE PREDICCIÓN

4.1. Introducción

En el capítulo anterior vimos la especificación inicial del modelo de predicción de obras de arte de autor, el cual consta de dos partes, una regresión hedónica del precio de las obras de arte de autor a partir de un conjunto de variables explicativas de dicho precio, combinado con el precio de venta pasado en aquellas ocasiones en las que se disponga información de una venta anterior de dicha obra artística.

Una vez especificado el modelo de predicción de obras de arte se procede con su validación mediante el análisis de las diferentes variables explicativas con el objetivo de seleccionar el modelo de predicción más adecuado, el cálculo de los parámetros de la regresión hedónica, el análisis del cumplimiento de las hipótesis básicas con el fin de determinar que no existen sesgos ni incongruencias en el modelo y finalmente, se destina una parte de los datos recopilados para la predicción y comprobación del funcionamiento del modelo en cuestión.

4.2. Estimación y análisis de los modelos de predicción

En este capítulo, se procede con el análisis del modelo especificado en el capítulo anterior y el cálculo de los parámetros de la regresión hedónica. En un primer lugar, comenzamos explicando el concepto de la regresión, los parámetros de la regresión y posteriormente estimaremos su valor e interpretaremos su significado.

El concepto de regresión hace referencia al modelo matemático que nos permite determinar⁶⁷ el valor de una variable (denominada como la variable dependiente o endógena) mediante otra variable o conjunto de variables (denominadas como las variables independientes o exógenas) a través de la relación de dependencia existente entre la variable endógena y el conjunto de variables exógenas. La regresión nos permite obtener un valor estimado que hace referencia a la esperanza matemática o media. La regresión es la función que asocia las medias condicionadas y puede ser lineal o no lineal, no obstante, se suele ajustar a la condición de linealidad ya que, por un lado, facilita en gran medida los cálculos y, por otro lado, es posible realizar transformaciones⁶⁸ en las variables independientes que nos permitan expresar de forma lineal relaciones no lineales entre variables.

La regresión hedónica o teoría de la demanda hedónica es un concepto que proviene del estudio de la economía y consiste⁶⁹ en la descomposición del artículo que es investigado en sus características constitutivas, obteniendo estimaciones del valor contributivo de cada característica.

La recta de regresión es la representación gráfica del modelo matemático mencionado, es decir, es la recta que une las medias condicionadas de la variable endógena a los valores de la variable exógena. Existen dos rectas de regresión. Por un lado, encontramos la recta de regresión real, la cual es desconocida al no conocer el valor real de los parámetros, y la recta de regresión estimada que es la que calculamos con el método que mencionamos posteriormente. A continuación se presenta el modelo

⁶⁷ Martínez de Ibarreta et al. (2019)

⁶⁸ Aplicar logaritmos a las variables, elevar al cuadrado o introducir variables de interacción como la multiplicación entre dos variables.

⁶⁹ Larrosa (2021)

matemático⁷⁰ de una regresión lineal simple y su representación gráfica en la que cada punto constituye una observación en concreto.

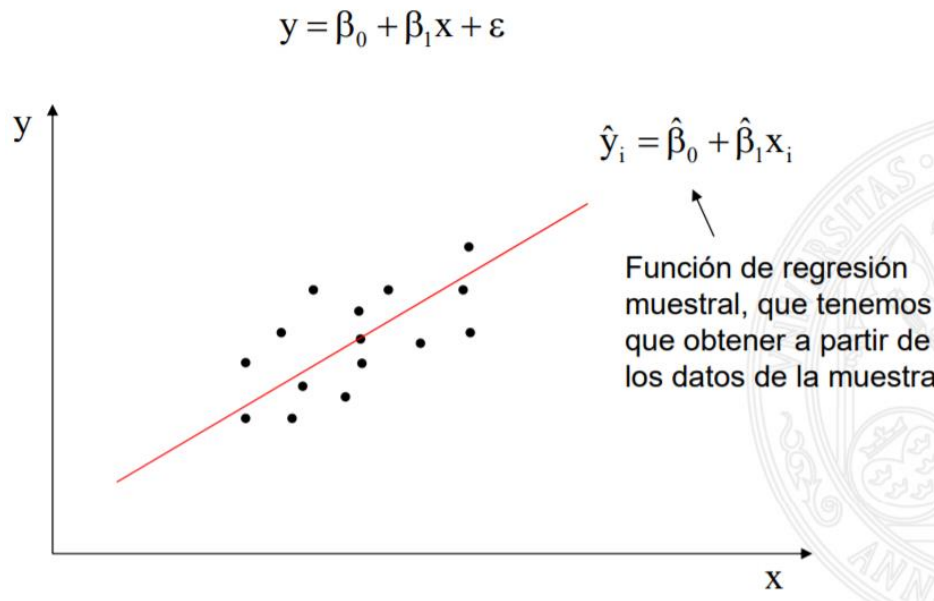


Ilustración 25. Fuente: Alvarez et al. (2021)

Los parámetros de la regresión miden el impacto de las variables explicativas en la variable endógena o dependiente, son cantidades que se suponen constantes y son siempre desconocidos. El valor real de los parámetros suele denotarse mediante la letra β , mientras que las estimaciones de dichos parámetros se denotan mediante la letra b .

El primer parámetro de todos (β_1) es el termino independiente⁷¹ y no se encuentra asociado a ninguna variable en especial. Su inclusión en el modelo es obligatoria ya que se trata de un término de ajuste, no obstante, su valor carece de significado (en

⁷⁰ Para más información consultar el capítulo anterior en el que se desarrolla el modelo matemático con mayor detalle.

⁷¹ Conocido como el intercepto.

determinadas situaciones podría interpretarse como el valor de la variable dependiente o endógena cuando todas las variables independientes o exógenas son cero, pero generalmente que todas las variables tomen el valor de cero no tiene sentido alguno).

El resto de los parámetros se encuentran ligados a una variable y su significado puede variar en función de la especificación de la variable a la que se encuentran asociados. En nuestro modelo existen dos interpretaciones para este tipo de parámetros:

- Los parámetros asociados a variables cuantitativas continuas o discretas representan el efecto marginal, es decir, el cambio en términos absolutos que experimenta la variable endógena cuando la variable exógena se incrementa en una unidad. El signo del parámetro nos indica si dicho cambio es positivo o negativo en la variable endógena. En este caso la interpretación de β es la siguiente: $\beta_i = \frac{\Delta y}{\Delta x_i}$ si $\Delta x_2 = \Delta x_3 = \dots = \Delta x_k = 0$
- Por otro lado, los parámetros asociados a variables nominales tienen un significado diferente. Cabe recordar que este tipo de variables se introducían en el modelo como una serie de variables nominales dicotómicas con respecto a una categoría base⁷². Dichas variables toman el valor 1 si se encuentran en la categoría de la variable y toman el valor 0 en todos los casos complementarios. Por ello, los parámetros hacen referencia al efecto diferencial medio, es decir, la diferencia que experimenta la variable endógena al producirse un cambio de la categoría codificada como 0 (categoría base) a la categoría codificada como 1 (categoría que representa esta variable nominal dicotómica).

⁷² La categoría base era la opción más repetida dentro de la variable observada.

El último elemento que contienen todos los modelos basados en una regresión son las perturbaciones aleatorias o errores, las cuales representan⁷³ todo aquello que pueda influir en la variable endógena pero que no está incluido en el modelo, ya que es imposible incluir todas las variables que tienen un efecto en la variable endógena. Es importante no confundir este concepto con los residuos, los cuales representan la diferencia entre la predicción y la observación. A continuación se diferencian estos dos conceptos de manera gráfica y se presentan la recta de regresión real y estimada mencionadas anteriormente:

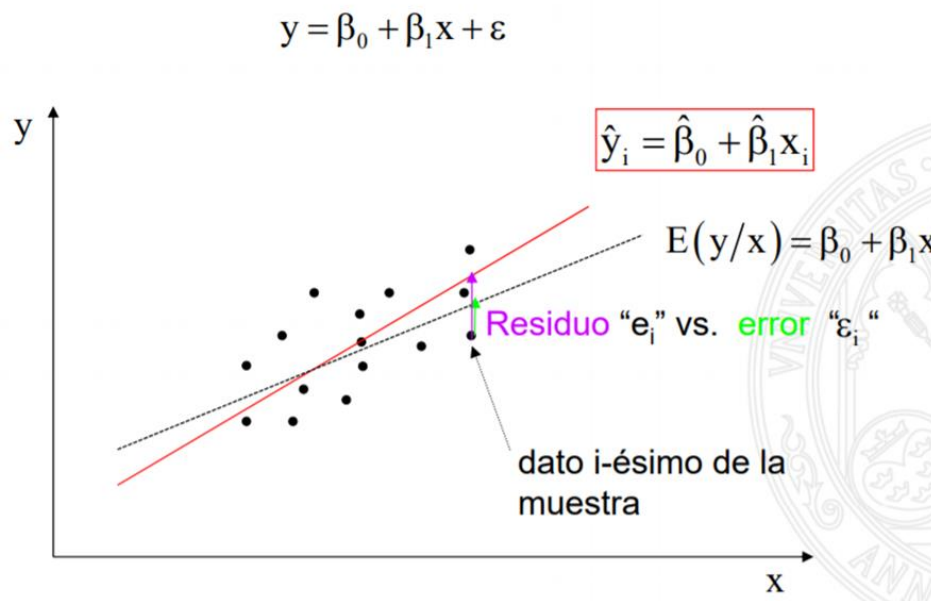


Ilustración 26. Fuente: Álvarez et al. (2021)

Como hemos visto, el valor real de los parámetros de la regresión es desconocido por lo que debemos calcular una estimación de los mismos. Los métodos más comunes para estimar estos parámetros son conocidos como el método de máxima verosimilitud y el método de los Mínimos Cuadrados Ordinarios (MCO). Para usar el método de máxima

⁷³ Martínez de Ibarreta et al. (2019)

verosimilitud debemos suponer una distribución de probabilidad para las perturbaciones aleatorias mientras que para el método MCO no, por lo que usaremos este último. El objetivo del método MCO es hacer mínimos los residuos⁷⁴, es decir, determinar el valor de los parámetros de tal forma que la suma de los cuadrados de los residuos sea la mínima posible. Se toma el cuadrado de los residuos ya que en caso contrario, los residuos positivos y negativos se contrarrestarían entre sí y el método perdería su eficacia. Tratamos de minimizar la siguiente función matemática:

$$S = \sum_{i=1}^n e_i^2 \quad \text{siendo} \quad e_i = y_i - \tilde{y}_i \quad [7]$$

Donde:

- $i \rightarrow$ Una observación concreta.
- $e_i \rightarrow$ Error o residuo
- $y_i \rightarrow$ Variable dependiente observada
- $\tilde{y}_i \rightarrow$ Variable dependiente estimada a través del modelo

Siguiendo esta metodología y cumpliendo una serie de hipótesis básicas, obtenemos unas estimaciones de los parámetros que cumplen unas propiedades⁷⁵ que convierten al modelo de los Mínimos Cuadrados Ordinarios como el método más preciso.

Para calcular las estimaciones de los parámetros debemos igual a cero la derivada de la función objeto (la función que se busca minimizar), la cual se presenta a continuación, respecto al vector de estimaciones de los parámetros.

⁷⁴ El residuo es la diferencia entre la predicción y la observación.

⁷⁵ Veremos estas propiedades más adelante en este capítulo.

$$\text{Min } F = \text{Min } \sum_{i=1}^n (y_i - \tilde{y}_i) \quad [8]$$

$$\frac{dF}{dB} = 2 * X^T * (Y - X * B) = 0 \quad [9]$$

De esta manera obtenemos la fórmula matemática para la estimación de los parámetros a través del método MCO. El vector de estimadores se calcula de la siguiente forma:

$$B = (X' * X)^{-1} * X' * Y \quad [10]$$

Donde:

- $Y \rightarrow$ Vector de observaciones de la variable dependiente
- $X \rightarrow$ matriz de variables independientes
- $B \rightarrow$ vector de estimadores
- $y_i \rightarrow$ Variable dependiente observada
- $\tilde{y}_i \rightarrow$ Variable dependiente estimada a través del modelo

Un último concepto importante para la comprensión de los resultados de la estimación de los parámetros de la regresión es el estadístico de contraste y el p-valor. El

estadístico de contraste o estadístico t es una variable aleatoria⁷⁶ que depende de la muestra recogida. Según el valor que tome, se compara con unas tablas base y se decide si aceptar o rechazar la hipótesis de que la variable exógena tiene influencia en el valor de la variable endógena. No obstante, el estadístico más importante es el p -valor que obtenemos a partir del estadístico t . El p -valor indica la probabilidad de rechazar una hipótesis cierta, por lo que solo en aquellos casos en los que encontremos una significación elevada⁷⁷ (determinada por un p -valor bajo) podremos concluir que la variable independiente tiene influencia en la determinación de la variable dependiente.

La estimación de los parámetros en nuestro modelo se lleva a cabo a través de la aplicación GRETL siguiendo la metodología de los Mínimos Cuadrados Ordinarios. Para ello se utilizan las 153 observaciones recogidas en la base de datos especificada en el capítulo anterior. A continuación se presentan los resultados obtenidos.

⁷⁶ Carollo (2011-2012)

⁷⁷ La significación elevada generalmente se relaciona con una probabilidad de rechazar una hipótesis cierta del 10%, 5% o 1%.

Modelo 7: MCO, usando las observaciones 1-153
Variable dependiente: Precio

	coeficiente	Desv. típica	Estadístico t	valor p	
const	12454.6	2299.24	5.417	2.75e-07	***
Tamano	-0.116475	0.197671	-0.5892	0.5567	
Siglo	-595.714	116.159	-5.128	1.01e-06	***
Paleta_Colores	-158.061	249.271	-0.6341	0.5271	
Nitidez	117.962	254.424	0.4636	0.6437	
Suavidad	721.633	252.515	2.858	0.0050	***
Claridad	-54.1266	275.699	-0.1963	0.8447	
Origen_Europa	755.350	272.917	2.768	0.0065	***
Origen_Asia	910.227	409.182	2.225	0.0278	**
Origen_Africa	221.839	719.580	0.3083	0.7583	
Casa_Subastas_Ja~	-570.509	270.639	-2.108	0.0369	**
Casa_Subastas_Ar~	-921.006	628.044	-1.466	0.1449	
Casa_Subastas_Ro~	-571.220	444.077	-1.286	0.2006	
Casa_Subastas_Hu~	-1148.09	493.343	-2.327	0.0215	**
Material_Papel	462.227	487.852	0.9475	0.3451	
Material_Acuarela	-209.834	520.818	-0.4029	0.6877	
Estilo_Paisaje	-55.2801	294.035	-0.1880	0.8512	
Estilo_Abstracta	-548.665	512.294	-1.071	0.2861	
Estilo_Surrealis~	-35.7187	373.652	-0.09559	0.9240	
Interaccion	0.185330	0.393259	0.4713	0.6382	

Ilustración 27. Fuente: Elaboración Propia

Analizando los resultados expuestos podemos observar cómo existen 6 variables cuya influencia es significativa en el cálculo del precio de las obras de arte con una significación superior al 95%. En cambio, solo existen 3 variables influyentes con un nivel de significación superior al 99%, es decir, una probabilidad de cometer error menor del 1%. Las variables más significativas en la determinación del precio de las obras de autor son el siglo de su creación, la suavidad, el origen y la casa de subastas. Dichas variables son las más determinantes en el precio de venta de las obras de arte.

La variable siglo de creación tiene un parámetro asociado cuyo valor es -596, lo cual nos indica que existe una relación inversa entre el precio de las obras artísticas y su siglo de creación. Las obras más antiguas suelen venderse por un precio mayor.

Por otro lado, podemos comprobar como la variable suavidad tiene un p-valor asociado del 5% y el parámetro asociado a dicha variable tiene un valor de 722, lo cual nos permite determinar con un nivel de significación del 95% que las obras artísticas que tienden a ser más suaves en cuanto a contornos y formas, alejándose de los vértices y las formas puntiagudas, tienen un precio medio 722 € mayor que otra obra artística con las mismas características pero una tendencia mayor a las vértices y formas puntiagudas.

Adicionalmente, otra de las variables cuya significación es superior al 90% es el origen o procedencia de la obra en cuestión. Todos los parámetros vinculados a las variables que analizan el origen de la obra son positivos, lo cual nos indica que las obras artísticas de procedencia americana (categoría base) tienen un precio inferior a los orígenes europeo y asiático. Aunque esta relación también se mantiene para aquellas obras de origen africano, su significación es muy baja por lo que no podemos concluir que esto se mantenga.

Finalmente, la casa de subastas en la que se vende la obra también juega un papel fundamental en la determinación del precio del final de venta. Todos los parámetros asociados a las variables que analizan la casa de subastas en la que se vendió la obra tienen carácter negativo, lo que nos indica que la casa de subastas Clars Auction Gallery tiene un precio de venta medio mayor. No obstante, solo los parámetros asociados a las casas de subastas Jasper 52 y Royal Treasures Warehouse LLC tienen un p-valor inferior al 10%. por ello, solo podemos concluir dicha afirmación para estas dos casas de subastas.

El análisis del modelo nos ofrece otros indicadores que nos aportan más información:

Media de la vble. dep.	1121.967	D.T. de la vble. dep.	1598.713
Suma de cuad. residuos	2.05e+08	D.T. de la regresión	1241.791
R-cuadrado	0.472085	R-cuadrado corregido	0.396668
F(19, 133)	6.259702	Valor p (de F)	3.27e-11
Log-verosimilitud	-1296.400	Criterio de Akaike	2632.801
Criterio de Schwarz	2693.409	Crit. de Hannan-Quinn	2657.421

Ilustración 28. Fuente: Elaboración Propia (2021)

En primer lugar, procedemos a explicar los diferentes indicativos estadísticos que acabamos de presentar. El R-Cuadrado o bondad de ajuste queda definido como la proporción de la variabilidad de la variable endógena explicada⁷⁸ por la variabilidad de las variables exógenas. En nuestro caso, corresponde con el grado en que los distintos precios de las obras de arte quedan explicados por las variables incluidas en el modelo. La fórmula matemática para su cálculo es la siguiente:

$$R^2 = \frac{Cov(x,y)^2}{V(x)*V(y)} \quad [11]$$

Donde:

x → Variable independiente

y → Variable dependiente

Para la comparación entre modelos se emplea el estadístico R-Cuadrado corregido ya que R-Cuadrado depende del número de variables introducidas en el modelo, por lo que un modelo con más variables puede tener un R-Cuadrado mayor que otro modelo con menor número de variables aunque sea peor modelo. A ello se le añade el estadístico F y su p-valor que nos indican si el modelo es significativo en su conjunto, es decir, nos indica si el R-Cuadrado es estadísticamente significativo.

⁷⁸ Martínez de Ibarreta et al. (2019)

Como podemos comprobar nuestro modelo posee un R-Cuadrado de 0.472 y un p-valor del estadístico F muy pequeño. Por ello se trata de un modelo que es significativo y que la variabilidad de la variable dependiente queda explicada en 47.2% por la variabilidad de las variables independientes.

Adicionalmente, aparecen otros estadísticos como el criterio de Akaike, el criterio de Schwarz y el criterio de Hannan-Quinn. Todos ellos son criterios de información con los que se calcula una medida de como de próximo está el modelo formulado al verdadero modelo generador de los datos. Por ello, siempre que se compare entre modelos se debe escoger aquel que tiene unos criterios de información más pequeños.

A la hora de especificar un modelo existen dos objetivos fundamentales. Por un lado, podemos estar buscando las causas que determinan el valor de la variable dependiente o por otro lado, podemos estar especificando el modelo con el objetivo de realizar predicciones en el futuro. Si nuestro objetivo es saber las causas, debemos buscar un modelo con el menor número de variables posibles para poder estimar los parámetros de la mejor forma posible. Si nuestro objetivo es la predicción, nos debemos enfocar en aumentar el R-Cuadrado aunque nuestro modelo contenga un número de variables demasiado elevado. A continuación, analizamos, calculamos los parámetros de un modelo en el que únicamente se incluyen las variables más significativas con el objetivo de comparar los resultados con el modelo descrito anteriormente y analizar si se mantienen las tendencias identificadas. Los resultados de este nuevo modelo se presentan a continuación:

Modelo 8: MCO, usando las observaciones 1-153
Variable dependiente: Precio

	coeficiente	Desv. típica	Estadístico t	valor p	
const	12606.2	1941.28	6.494	1.29e-09	***
Siglo	-611.471	99.1231	-6.169	6.68e-09	***
Suavidad	696.095	217.568	3.199	0.0017	***
Origen_Europa	738.567	257.276	2.871	0.0047	***
Origen_Asia	726.738	383.543	1.895	0.0601	*
Origen_Africa	45.2402	582.088	0.07772	0.9382	
Casa_Subastas_Ja~	-481.938	242.904	-1.984	0.0492	**
Casa_Subastas_Ar~	-634.741	416.258	-1.525	0.1295	
Casa_Subastas_Ro~	-420.163	410.868	-1.023	0.3082	
Casa_Subastas_Hu~	-1049.14	470.519	-2.230	0.0273	**
Media de la vble. dep.	1121.967	D.T. de la vble. dep.	1598.713		
Suma de cuad. residuos	2.13e+08	D.T. de la regresión	1220.820		
R-cuadrado	0.451401	R-cuadrado corregido	0.416874		
F(9, 143)	13.07380	Valor p (de F)	4.41e-15		
Log-verosimilitud	-1299.340	Criterio de Akaike	2618.681		
Criterio de Schwarz	2648.985	Crit. de Hannan-Quinn	2630.991		

Ilustración 29. Fuente: Elaboración Propia (2021)

En este nuevo modelo podemos comprobar como todas las relaciones identificadas anteriormente se mantienen y en el algún caso su p-valor disminuye, lo que indica que dichas relaciones se hacen más significativas. El indicador R-Cuadrado disminuye, lo que indica que en este modelo la variabilidad de la variable endógena queda determinada en menor medida por la variabilidad de las variables exógenas. No obstante, el R-Cuadrado corregido aumenta por lo que es un mejor modelo en términos de precisión de los parámetros.

El objetivo de la selección del modelo predictivo es escoger aquel que sea el más apropiado y tratar de minimizar el número de variables. Por ello, emplearemos el último modelo que ha sido optimizado.

Como hemos comprobado anteriormente, no se ha identificado una relación estable entre el tamaño de la obra artística y su precio de venta, no obstante, creemos que el tamaño de la obra tiene alta probabilidad de influir en el valor de la obra. Por ello hemos comprobado esta relación manteniendo todas las demás variables constantes. Este grafico se muestra a continuación:

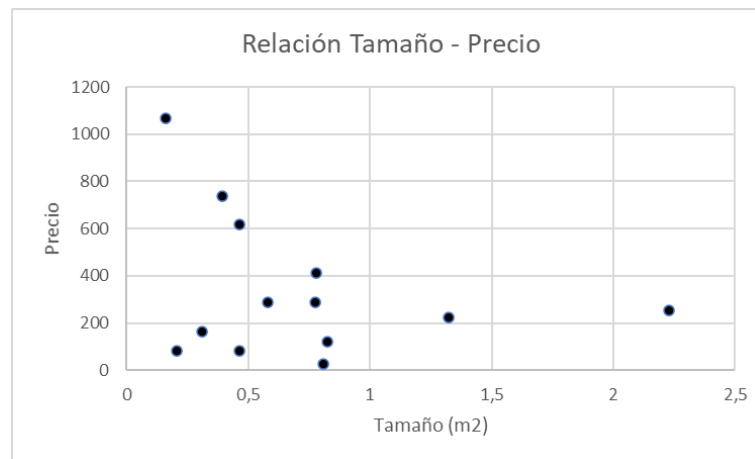


Ilustración 30. Fuente: Elaboración Propia (2021)

Se puede apreciar cómo puede existir un máximo en el precio para un tamaño cercano al DIN A1 y para tamaños superiores el precio de venta disminuye. No obstante, no se puede concluir que exista una relación ya que el número de observaciones con las que desarrollamos esta conclusión no es suficientemente significativo, por lo que no incluiremos esta variable en el modelo final al no tener suficiente significación estadística en nuestra base de datos.

4.3. Estimadores de los parámetros de la regresión

Como hemos comentado en el capítulo anterior, si se cumplen ciertas hipótesis básicas⁷⁹, los estimadores MCO obtienen unas propiedades deseables que convierten al método MCO en uno de los más útiles. En este capítulo vamos a analizar que se cumplen estas hipótesis. A continuación se presentan dichas hipótesis:

- **Media nula:** Esta hipótesis supone que el valor esperado de las perturbaciones aleatorias es cero. Esta hipótesis se cumple siempre que se introduzca el termino independiente en la formulación del modelo.
- **Homocedasticidad:** Esta hipótesis supone que la varianza de las perturbaciones aleatorias es constante e igual para todas las observaciones.

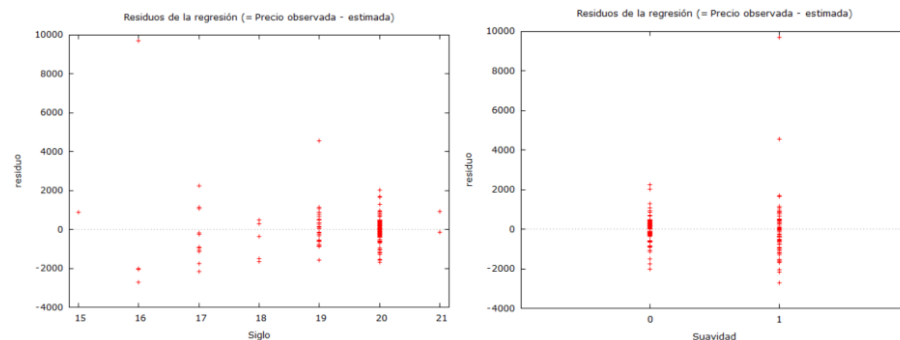


Ilustración 31. Fuente: Elaboración Propia (2021)

En estos gráficos se puede apreciar cómo no existen una constancia en la varianza de las perturbaciones aleatorias frente a las variables siglo y suavidad. No obstante, para asegurar que se cumple esta hipótesis básica acudimos al contraste White. Este contraste nos permite comprobar si existe

⁷⁹ El análisis de las siguientes hipótesis se basa en el capítulo 3.3 del libro escrito por Martínez de Ibarreta et al. (2019)

homocedasticidad ya que unifica en una ecuación los cuadrados y los productos cruzados de todas las variables de la regresión. La hipótesis nula de este contraste consiste en la existencia de homocedasticidad, por lo que debemos acudir al p-valor de este contraste para comprobar si se cumple la hipótesis de heteroscedasticidad. En la tabla siguiente se presentan los resultados.

```
Contraste de heterocedasticidad de White -  
Hipótesis nula: [No hay heterocedasticidad]  
Estadístico de contraste: LM = 31.814  
con valor p = P(Chi-cuadrado(29) > 31.814) = 0.328096
```

Ilustración 32. Fuente: Elaboración Propia (2021)

El p-valor es 0.328 por lo que no se puede rechazar la hipótesis nula y se cumple la segunda condición de homocedasticidad.

- **No existe autocorrelación:** Esta hipótesis supone que las perturbaciones aleatorias de diferentes observaciones no tienen relación lineal entre sí, es decir, no existen ningún tipo de relación entre lo no explicado para diferentes observaciones. Acudimos a la matriz de correlaciones para comprobar si existe alguna correlación entre variables. A continuación se presenta dicha matriz:

```

      Siglo      Suavidad      Origen_Europa      Origen_Asia
1.0000      -0.1412      -0.3870      0.1739 Siglo
      1.0000      0.2860      -0.1347 Suavidad
      1.0000      -0.3241 Origen_Europa
      1.0000 Origen_Asia

Origen_Africa Casa_Subastas_~ Casa_Subastas_~ Casa_Subastas_~
0.0717      0.0503      0.1586      -0.0548 Siglo
0.0501      -0.1478      0.0454      -0.1039 Suavidad
-0.1416      0.1053      -0.2145      -0.0866 Origen_Europa
-0.0773      -0.1350      -0.0463      0.3941 Origen_Asia
1.0000      0.0865      -0.0512      0.0573 Origen_Africa
      1.0000      -0.2145      -0.2633 Casa_Subastas_~
      1.0000      -0.0951 Casa_Subastas_~
      1.0000 Casa_Subastas_~

Casa_Subastas_~
0.1443 Siglo
-0.0765 Suavidad
-0.0237 Origen_Europa
0.4215 Origen_Asia
-0.0536 Origen_Africa
-0.2248 Casa_Subastas_~
-0.0812 Casa_Subastas_~
-0.0997 Casa_Subastas_~
1.0000 Casa_Subastas_~

```

Ilustración 33. Fuente: Elaboración Propia (2021)

Como se puede comprobar ninguna correlación entre variables alcanza un valor elevado⁸⁰, por lo que podemos concluir que también se cumple esta hipótesis sobre la ausencia de correlación.

- **Ausencia de colinealidad perfecta:** Esta hipótesis impide la existencia de una relación lineal exacta entre dos variables explicativas del modelo de predicción. Si no se cumple esta hipótesis no es posible la estimación del modelo ni la estimación de los parámetros. Este defecto es fácil de identificar ya que ocurre en aquellos casos en los que el rango de la matriz de las variables independientes no es máximo. Nuestra matriz de las variables independientes tiene un rango de 153 y ninguna fila es combinación lineal de otras filas, por lo que no nos encontramos antes esta problemática.

⁸⁰ Un valor superior a 0.8 genera indicios de la existencia de autocorrelación.

Tras el cumplimiento de las hipótesis estudiados podemos concluir que nuestros estimadores de los parámetros pueden ser denominados estimadores MCO y por ello cumplen una serie de propiedades⁸¹:

- **Insesgados:** la esperanza matemática de los estimadores coincide con el valor real de los parámetros, es decir, $E[b_i] = \beta_i$.
- **Óptimos:** los estimadores seleccionados son los que menor varianza tienen de todos los estimadores lineales e inseguros existentes.
- **Consistentes:** Si el tamaño de la muestra tiende a infinito, la probabilidad de que exista distancia entre el verdadero parámetro y la estimación tiende a cero.

4.4. Inclusión de la información de ventas pasadas en el modelo

Como comentamos en el capítulo anterior, en aquellos casos en los que disponemos de la información de obras que han sido vendidas en el pasado, introducimos dicha información en el modelo con el objetivo de obtener una predicción más precisa y con un menor error. No obstante, antes de combinar ambas predicciones especificamos un modelo que nos permita ajustar el precio de venta pasado al precio de venta actual con una regresión lineal en la que incorporamos el número de meses transcurridos desde la venta pasada hasta la actualidad. Este modelo fue presentado en el capítulo anterior, no obstante, se recuerda:

$$\omega_{2,j} = \beta_{0r} + \beta_{1r} * \omega_{p,i} + \beta_{2r} * R_j \quad [12]$$

Donde:

⁸¹ Martínez de Ibarreta et al. (2019)

$\omega_{1,j}$	Precio de venta (Regresión hedónica)	β_n	Parámetro de la regresión múltiple	R_j	N.º de Meses Transcurridos entre las ventas
$\omega_{2,j}$	Precio de venta (Regresión con venta pasada)	$\omega_{p,i}$	Precio de venta (Venta pasada)	α	Parámetro del modelo combinado
$\omega_{3,j}$	Precio de venta (Modelo combinado)				

Tabla 3. Fuente: Elaboración Propia (2021)

A continuación, procedemos con el estudio del modelo presentado a través del cálculo de los parámetros y los indicadores estadísticos necesarios. Para la estimación de los parámetros empleamos de nuevo el método de los Mínimos Cuadrados Ordinarios por todas las ventajas que nos ofrece (mencionadas anteriormente). Los resultados obtenidos son:

Modelo 1: MCO, usando las observaciones 1-15

Variable dependiente: Precio_Venta

	coeficiente	Desv. típica	Estadístico t	valor p	
const	275.361	338.202	0.8142	0.4314	
Precio_Venta_Pas~	0.799424	0.225191	3.550	0.0040	***
N_Meses_Transcur~	-11.8583	45.7646	-0.2591	0.7999	
Media de la vble. dep.	1238.133	D.T. de la vble. dep.	870.8529		
Suma de cuad. residuos	4894584	D.T. de la regresión	638.6564		
R-cuadrado	0.539003	R-cuadrado corregido	0.462170		
F(2, 12)	7.015270	Valor p (de F)	0.009598		
Log-verosimilitud	-116.5010	Criterio de Akaike	239.0020		
Criterio de Schwarz	241.1262	Crit. de Hannan-Quinn	238.9794		

Ilustración 34. Fuente: Elaboración Propia (2021)

La variable Numero de Meses Transcurridos tiene un p-valor cercano a 0.8 lo que nos indica que su significación es muy baja y la influencia de esta variable en el precio de venta actual es reducida. Por ello, concluimos que no es efectivo ajustar el precio de venta actual partiendo del precio de venta en el pasado a través del número de meses transcurridos entre ambas fechas, por lo que eliminamos esta variable del modelo. A continuación se presentan los nuevos resultados:

Modelo 2: MCO, usando las observaciones 1-15
Variable dependiente: Precio_Venta

	coeficiente	Desv. típica	Estadístico t	valor p	
const	242.399	301.913	0.8029	0.4365	
Precio_Venta_Pas~	0.777027	0.200342	3.879	0.0019	***
Media de la vble. dep.	1238.133	D.T. de la vble. dep.	870.8529		
Suma de cuad. residuos	4921969	D.T. de la regresión	615.3154		
R-cuadrado	0.536424	R-cuadrado corregido	0.500764		
F(1, 13)	15.04285	Valor p (de F)	0.001902		
Log-verosimilitud	-116.5428	Criterio de Akaike	237.0857		
Criterio de Schwarz	238.5018	Crit. de Hannan-Quinn	237.0706		

Ilustración 35. Fuente: Elaboración Propia (2021)

Este es el modelo que emplearemos para la estimación del precio de venta actual a partir del precio de venta en el pasado, ya que tiene un R-Cuadrado Corregido mayor que el modelo anterior, los tres criterios de información presentados son menores que en el modelo anterior.

El modelo final para la predicción del valor de obras de arte de autores sin gran renombre se presenta a continuación:

$$\omega_{3,j} = \alpha * \omega_{1,j} + (1 - \alpha) * \omega_{2,j} \quad [13]$$

$$\omega_{1,j} = \beta_0 + \beta_1 * S_j + \sum_{i=2}^{i=4} (\beta_i * P_{i,j}) + \sum_{x=5}^{x=8} (\beta_x * C_{x,j}) + \beta_9 * F_j + \mu_j \quad [14]$$

$$\omega_{2,j} = \beta_{0'} + \beta_{1'} * \omega_{p,i} + \mu_j \quad [15]$$

Donde:

$\omega_{1,j}$	Precio de venta (Regresión hedónica)	β_n	Parámetro de la regresión múltiple	$P_{i,j}$	Origen	i=3 → Europa i=4 → Asia i=5 → África
$\omega_{2,j}$	Precio de venta (Regresión con venta pasada)	$\omega_{p,i}$	Precio de venta (Venta pasada)	$C_{x,j}$	Casa de subastas	x=6 → Jasper52 x=7 → ArtForSale.com x=8 → Royal Treasures Warehouse LLC x=9 → Hudson Valley Auctions
$\omega_{3,j}$	Precio de venta (Modelo combinado)	S_j	Siglo	α	Parámetro del modelo combinado	
F_j	Suavidad					

Tabla 4. Fuente: Elaboración Propia (2021)

4.5. Predicción con el modelo predictivo

Una vez seleccionado el modelo final, procedemos con su utilización para la predicción, para ello se utilizan desde la observación 15 a la 50. Calcularemos el valor de la predicción con el modelo en base a las variables identificadas y los parámetros calculados y posteriormente compararemos dichos resultados con el valor real del precio de venta de la obra.

En primer lugar, se comprueba la capacidad predictiva del modelo basado únicamente en la regresión basada en los atributos de las obras artísticas, es decir, sin la información de aquellas obras que ya han sido vendidas en el pasado. Posteriormente, analizaremos la predicción una vez se incluye esta información comprobando si se mejoran las estimaciones o por el contrario disminuye su precisión. A continuación se presenta una gráfica con las predicciones del modelo y el precio real de venta, junto con un intervalo del 95% respecto a la predicción del precio de venta.

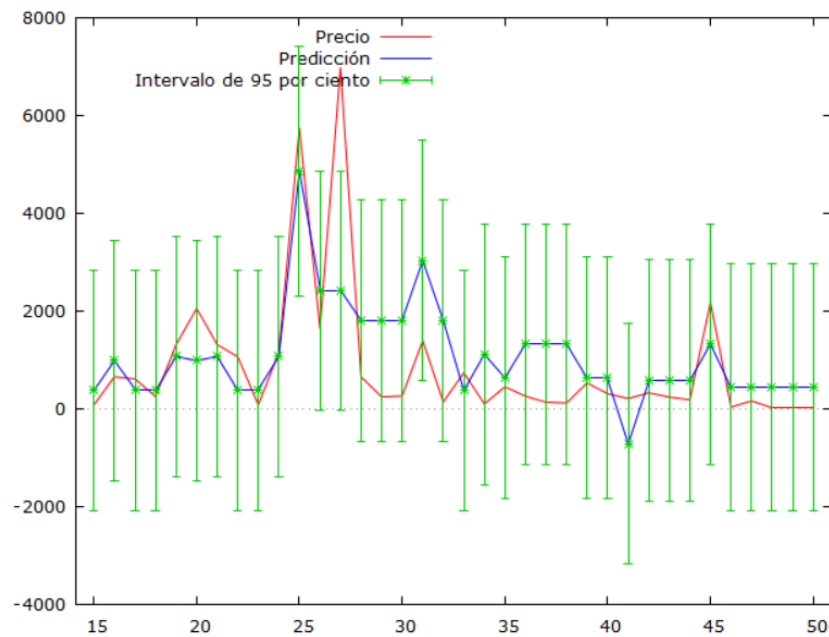


Ilustración 36. Fuente: Elaboración Propia (2021)

Adicionalmente, se presentan los resultados obtenidos para la realización de dicha gráfica:

Para intervalos de confianza 95%, $t(143, .0.025) = 1.977$

	Precio	Predicción	Desv. Típica	Intervalo de confianza :
15	82.00	376.80	1240.521	-2075.32 - 2828.93
16	658.00	988.27	1240.247	-1463.31 - 3439.86
17	616.00	376.80	1240.521	-2075.32 - 2828.93
18	255.00	376.80	1240.521	-2075.32 - 2828.93
19	1315.00	1072.90	1242.513	-1383.17 - 3528.96
20	2055.00	988.27	1240.247	-1463.31 - 3439.86
21	1315.00	1072.90	1242.513	-1383.17 - 3528.96
22	1068.00	376.80	1240.521	-2075.32 - 2828.93
23	82.00	376.80	1240.521	-2075.32 - 2828.93
24	1151.00	1072.90	1242.513	-1383.17 - 3528.96
25	5753.00	4868.82	1289.871	2319.14 - 7418.50
26	1644.00	2422.94	1239.680	-27.53 - 4873.40
27	6986.00	2422.94	1239.680	-27.53 - 4873.40
28	658.00	1811.47	1246.673	-652.82 - 4275.76
29	247.00	1811.47	1246.673	-652.82 - 4275.76
30	267.00	1811.47	1246.673	-652.82 - 4275.76
31	1397.00	3034.41	1240.593	582.14 - 5486.68
32	144.00	1811.47	1246.673	-652.82 - 4275.76
33	740.00	376.80	1240.521	-2075.32 - 2828.93
34	103.00	1118.14	1351.075	-1552.52 - 3788.80
35	452.00	633.43	1247.012	-1831.53 - 3098.39
36	263.00	1329.53	1245.675	-1132.79 - 3791.84
37	140.00	1329.53	1245.675	-1132.79 - 3791.84
38	123.00	1329.53	1245.675	-1132.79 - 3791.84
39	534.00	633.43	1247.012	-1831.53 - 3098.39
40	312.00	633.43	1247.012	-1831.53 - 3098.39
41	214.00	-716.60	1244.459	-3176.52 - 1743.31
42	329.00	590.96	1251.453	-1882.78 - 3064.70
43	244.00	590.96	1251.453	-1882.78 - 3064.70
44	188.00	590.96	1251.453	-1882.78 - 3064.70
45	2180.00	1329.53	1245.675	-1132.79 - 3791.84
46	41.00	438.16	1279.675	-2091.37 - 2967.68
47	164.00	438.16	1279.675	-2091.37 - 2967.68
48	41.00	438.16	1279.675	-2091.37 - 2967.68
49	41.00	438.16	1279.675	-2091.37 - 2967.68
50	41.00	438.16	1279.675	-2091.37 - 2967.68

Estadísticos de evaluación de la predicción utilizando 36 observaciones

Error medio -199.77
Raíz del Error cuadrático medio 1104.2

Ilustración 37. Fuente: Elaboración Propia (2021)

Como se puede comprobar, solo un precio de venta observado se encuentra fuera del intervalo del 95% de la predicción, lo cual nos indica que en estas observaciones el modelo se adapta bien a las observaciones, no obstante, cuando aparecen precio muy

altos o bajos, conocidos como “outliers”, el modelo baja en gran medida su precisión al no ser capaz de captar estos grandes cambios.

A continuación, se presenta una tabla con el análisis de aquellas variables de las que se dispone de la información del precio de venta en el pasado. Se presenta tanto la estimación del modelo combinado en el que se incluye la información mencionada como la predicción basándonos únicamente en el modelo de regresión de los atributos de la obra artística. Adicionalmente, se presenta la información de la diferencia entre el precio observado y la predicción del modelo basado en la regresión de atributos de la obra (Dif 1) y la diferencia entre el precio observado y la predicción del modelo mejorado con la información de las ventas del pasado.

Predicción Modelo Ventas Pasadas	Predicción Modelo Carac	Predicción Final	Precio Observado	Dif 1	Dif Final
1686,115166	1072,9	1379,507583	1569	496,1	189,492417
1526,047604	2422,94	1974,493802	2271	-151,94	296,506198
1140,642212	1726,84	1433,741106	2065	338,16	631,258894
1461,554363	1072,9	1267,227182	1404	331,1	136,772819
1054,392215	1072,9	1063,646108	1074	1,1	10,3538925
2295,304334	621,6	1458,452167	1074	452,4	-384,452167
284,358458	1329,53	806,944229	54	-1275,53	-752,944229
1045,067891	-105,13	469,9689455	1177	1282,13	707,031055
1397,061122	633,43	1015,245561	496	-137,43	-519,245561
1846,959755	1941	1893,979878	2829	888	935,020123
1268,851667	1941	1604,925834	1404	-537	-200,925834
2167,094879	1941	2054,04744	2271	330	216,952561
836,047628	506,34	671,193814	826	319,66	154,806186
306,892241	683,38	495,1361205	41	-642,38	-454,136121
255,608459	-43	106,3042295	17	60	-89,3042295

Tabla 5. Fuente: Elaboración Propia (2021)

En estas obras en las que disponemos de la información del precio de venta en el pasado, las predicciones son más eficaces por lo que la adición de esta información al modelo nos permite mejorar los resultados en termino de precisión y disminuir el error. Esto se puede comprobar en la tabla que se presenta a continuación en la que se incluye la suma al cuadrado de las dos diferencias mencionadas.

Suma Dif 1	Suma Dif Total
5691816,042	3225244,495

Tabla 6. Fuente: Elaboración Propia (2021)

En esta tabla se puede observar como la suma al cuadrado de las diferencias entre la predicción y el precio observado es menor en el modelo combinado en el que se utiliza toda la información disponible.

4.6. Conclusiones

En este capítulo hemos realizado el análisis de los modelos propuestos mediante su estudio y el cálculo de las estimaciones de los parámetros de la regresión con el objetivo de mejorarlos y seleccionar el modelo más adecuado, junto con la identificación de tendencias existentes en el mercado del arte y el posterior uso del modelo para la predicción.

En un primer lugar, hemos comprobado como el modelo de regresión más adecuado era aquel en el que se eliminaban una serie de variables y dejábamos solo aquellas que han resultado ser significativas a la hora de determinar el precio de las obras artísticas. Se ha

identificado que el precio está relacionado con el siglo de creación, la suavidad de la obra, su origen y la casa de subastas en la que fue vendida.

Posteriormente, se ha comprobado que los estimadores de los parámetros cumplían una serie de hipótesis básicas para asegurar que nos ofrecen una serie de propiedades y evitar que el modelo se encuentre sesgado.

También se ha analizado la inclusión en el modelo de la información procedente de aquellas obras que ya habían sido vendidas en el pasado, comprobando que esta información mejora la estimación del modelo.

Finalmente, hemos podido utilizar el modelo final para la predicción de un número concreto de observaciones y su comparación el valor real de venta de cara a analizar el funcionamiento de nuestro modelo de predicción de obras de arte de artistas desconocidos.

En el siguiente capítulo, procederemos con una definición clara de los resultados y conclusiones obtenidas en el estudio y su comparación con los estudios existentes hasta la actualidad de cara a identificar similitudes o diferencias entre este estudio y aquellos que también se centraron en la industria del arte.

5.RESULTADOS, CONCLUSIONES Y DESARROLLOS FUTUROS

5.1. Introducción

En este capítulo vamos a presentar los resultados y conclusiones extraídos del proyecto, compararemos dicha información con los estudios relacionados con la industria del arte existentes hasta la actualidad tratando de identificar semejanzas y discordancias. Finalizaremos el capítulo con una sección de desarrollos futuros en la que identificaremos posibles campos de investigación en los que sería muy interesante focalizar futuros estudios.

Al centrar nuestro estudio en aquellas obras que no están firmadas o cuyos autores no tienen fama ni influencia en el precio de venta, no disponemos de información de comparación por la inexistencia de proyectos que se enfoquen en esta área de la industria del arte. No obstante, compararemos los resultados obtenidos en el presente estudio con aquellos analizados en obras de autores con cierta fama con el objetivo de identificar semejanzas o discordancias entre ambos.

5.2. Resultados, conclusiones y comparación con otros estudios

Los principales objetivos de este proyecto consistían en la elaboración de un modelo de predicción del valor de obras de arte y la consiguiente identificación de tendencias en el mercado del arte. Posteriormente, el análisis del estado del arte nos permitió conocer la inexistencia de estudios que se enfocasen en aquellas obras que no están firmadas por

ningún autor o que su autor no es famoso. Todos los estudios analizados se centraban en obras de autores famosos como Pablo Picasso o Claude Monet. Lo que nos indica que los estudios existentes se enfocan en el arte especulativo, es decir, la compra de obras artísticas por razones ajenas a los criterios estéticos y únicamente por el pintor que realizó dichas obras. Por el contrario, nuestro estudio se aleja del arte especulativo y se centra en el valor intrínseco de las obras artísticas determinada por sus características visuales.

En los estudios analizados en el estado del arte se pudo comprobar como el autor jugaba un papel fundamental en la determinación del precio de venta de la obra, el cual quedaba determinado en gran medida por esta característica y después se ajustaba en función de los atributos de la obra como el año de creación o el origen. También vimos cómo el momento en la carrera del autor en el que se pintó dicha obra jugaba un papel fundamental.

Al centrar nuestro estudio en aquellas obras que no fueron creadas por autores con gran renombre, toda esta investigación existente no nos aportaba ningún valor ya que no podíamos hacer uso de ella. Por ello, nos encontramos antes algo novedoso y debíamos investigar que variables tienen influencia en la valoración de este tipo de obras.

En un primer lugar, se especificó un modelo inicial con una serie de variables las cuales consideramos podían tener influencia en el precio de las obras artísticas. Posteriormente, pudimos comprobar como de estas numerosas variables solo algunas eran influyentes en el precio de venta, con la consiguiente reducción del modelo a un menor número variables. A continuación se presentan los resultados y tendencias identificadas en nuestro estudio:

- En primer lugar, el principal resultado del proyecto ha sido la creación de un modelo de predicción de obras de arte en las que sus autores no tienen influencia en la determinación del precio o no se encuentran firmadas por ningún autor. Dicho modelo posee un coeficiente de predicción de 0.45, el cual aumenta cuando se dispone del precio de venta de la obra en una etapa anterior.

Pese a no utilizar la misma base de datos que en este proyecto, otros estudios muestran como la realización de un modelo centrado en un único artista alcanza coeficientes de predicción alrededor de 0.58 (Hosny, 2017), lo que nos permite confirmar que los modelos centrados en un único artista tienen una mayor capacidad predictivo, precisión y un menor error.

Por ello, al no centrarnos en un autor en concreto, el coeficiente de predicción de nuestro modelo se reduce, por lo que consideramos que la mejor metodología para la predicción del valor de obras de arte es la combinación del modelo presentado en este estudio con el conocimiento de un experto tasador que pueda utilizar dicha predicción para ajustarla en aquellos casos en los que el modelo no pueda captar nuevas tendencias o influencias en el precio.

- Las obras artísticas tienden a revalorizarse con el paso de los siglos. En concreto en nuestro proyecto se identificó que de media una obra artística aumenta su precio en torno a 600⁸² € por siglo. La mayor diferencia en el precio se observa entre aquellas obras creadas en el siglo 21 y las obras cuya creación data entre los siglos 15 y 17.

Este resultado coincide con lo analizado en otros proyectos que muestran que por norma general, si comparamos dos piezas similares en calidad, tamaño, estado de conservación, etc. la más antigua será más valiosa (Trigo, 2021).

⁸² Todas las diferencias medias están calculadas a partir de las estimaciones de los parámetros de la regresión.

- Otra de las tendencias identificadas relaciona la aparente suavidad de la obra en cuestión y su precio de venta. Aquellas obras designadas como “suaves” hacen referencia a pinturas en las que se evitan los vértices, figuras puntiagudas y se tiende a trazar formas y contornos más redondeados y que dan una mayor sensación de suavidad. Las obras artísticas que presentan esta característica de suavidad tienden a venderse por un precio mayor. La base de datos recogida para nuestro estudio nos muestra que la diferencia media entre las obras que presentan esta cualidad y aquellas que no es de alrededor de 700 €⁸³.

Pese a la utilización de una base de datos distinta a la nuestra, Hosny también analizó la suavidad de las obras mediante un algoritmo⁸⁴ utilizado comúnmente como un detector de esquinas. La conclusión obtenida coincide con la presentada en este estudio, las obras artísticas con un porcentaje bajo de esquinas son más propensas a alcanzar precios de venta elevados.

- Otro de los descubrimientos de este estudio relaciona al precio de venta de las obras con su origen, es decir, el lugar donde fueron pintadas. Las obras artísticas realizadas en América muestran menores precios que aquellas realizadas en Europa, Asia o África. No obstante, solo podemos concluir que las obras realizadas en Europa muestran de media un precio superior en 740 € frente a aquellas realizadas en América con un nivel de significación superior al 99%. Adicionalmente, las obras realizadas en Asia muestran de media un precio superior en 730 € frente a aquellas realizadas en América con un nivel de significación superior al 90%. La relación entre las obras creadas en África y América no se sostiene estadísticamente debido a su bajo nivel de significación.

⁸³ Estos son valores medios y la diferencia en el precio depende en gran medida del rango de precio de la obra en el que nos encontremos.

⁸⁴ Harris Corner Detection

Numerosos estudios han destacado la relación existente entre la procedencia de una obra y su precio de venta. No obstante, todos los estudios analizan la procedencia desde el punto de vista de las colecciones por las que ha pasado la obra y casas de subastas en las que se ha vendido. Todos estos estudios destacan la relación existente entre su procedencia y su precio, aquellas obras que han pasado por colecciones o familias famosas aumentan su precio a igualdad en otros atributos de las obras. No obstante, no podemos comprar nuestros resultados ya que el origen desde el punto de vista de los continentes en los que se creó la obra no ha sido analizado.

- La casa de subastas en la que se vendió la obra también tuvo un papel determinante en el precio de la obra. Las obras artísticas vendidas en la casa de subastas Clars Auction Gallery tienen un precio medio mayor que aquellas vendidas en las otras casas de subastas⁸⁵ analizadas. No obstante, esta relación solo es significativa estadísticamente para las casas de subastas Jasper 52 y Hudson Valley Auctions. En estas casas de subastas se obtuvo un precio medio menor en 480 €⁸⁶ en Jasper 52 y 1.050 €⁸⁷ en Hudson Valley Auctions. Es importante destacar que estos resultados no implican que una misma obra artística sea vendida a un mayor precio en Clars Auction Gallery ya que no se encontró información de obras artísticas que hayan sido vendidas en dos casas de las estudiadas de manera simultánea. Los resultados únicamente nos permiten determinar que las obras en Clars Auction Gallery tienen de media un precio mayor frente a estas dos casas de subastas mencionadas anteriormente. Hay que destacar que ningún estudio ha analizado la relación entre las 5 casas de subastas presentadas en este proyecto y el precio de venta de las obras artísticas, por lo

⁸⁵ Jasper 52, Art For Sale, Royal Treasures Warehouse LLC, Hudson Valley Auctions

⁸⁶ Nivel de significación alrededor del 95%.

⁸⁷ Nivel de significación superior al 95%.

que no es posible contrastar nuestros resultados con los obtenidos en otros proyectos.

- En el estudio también se analizaron otras variables: tamaño de la obra, paleta de colores⁸⁸, nitidez⁸⁹, claridad⁹⁰, material⁹¹, estilo⁹² y una variable de interacción entre las obras de estilo abstracto y su tamaño⁹³. De las tendencias identificadas respecto al precio de venta de la obra y estas variables estudiadas ninguna posee la suficiente significación para poder considerarse como una relación válida. Por ello, determinamos que todas estas variables tienen una menor relevancia en el precio frente a las identificadas anteriormente.
- El último resultado extraído de este estudio se obtuvo al analizar el modelo que relacionaba el precio de venta de aquellas obras de las que se disponía de su precio de venta en una fecha anterior. Se estudio si el espacio temporal⁹⁴ entre ambas fechas podía tener influencia en el nuevo precio de venta. No se obtuvo una relación clara entre el tiempo transcurrido y el nuevo precio de venta de la obra, lo que nos indica que así como el precio de venta cambia con los siglos, no se aprecia una relación clara cuando este espacio temporal es de varios meses o algunos años.

Pese a la creencia de que puede existir una relación entre el tamaño de la obra artística y su precio de venta, otros estudios obtuvieron los mismos resultados

⁸⁸ Esta variable es dicotómica y se diferencia entre aquellas obras realizadas con una variedad de colores y otras realizadas en un color con diferentes tonalidades.

⁸⁹ Obras con un gran detalle y otras que tienden a difuminarse y mezclar los contornos de las figuras representadas.

⁹⁰ Obras en colores claros y obras en colores más oscuros.

⁹¹ Si distinguían 3 materiales: lienzo, papel y acuarela.

⁹² Se distinguían 4 estilos: retratos/escenas, paisajes, abstracto y surrealista.

⁹³ Tratando de ajustar la diferencia media de tamaño existente entre las obras de estilo abstracto y las obras del resto de estilos.

⁹⁴ Medido en número de meses transcurridos.

que el presente proyecto. En la valoración de pinturas de la India no existe una relación clara entre el tamaño de la obra artística y su correspondiente precio de venta (Ghosh, Mehta, 2017). No obstante, es interesante seguir investigando esta relación como veremos a continuación.

5.3. **Desarrollos Futuros**

Este proyecto presenta numerosas áreas en las que continuar la investigación es muy interesante dada la falta de información. A continuación se presentan aquellos desarrollos futuros que consideramos pueden aportar resultados más fructíferos:

- Un primer campo de investigación que resulta interesante analizar es la paleta de colores. En este estudio se analizaron las posibles relaciones entre el precio de venta y la diferencia entre aquellas obras que presentan una variedad de colores y aquellas que se realizan en un color con diferentes tonalidades. Pese a no encontrar ninguna relación válida entre el precio de venta de la obra y la paleta de colores, creemos que sería muy interesante analizar la variable de los diferentes colores en mayor detalle incluyendo un estudio de ella mediante diferentes variables que nos permitan analizar cada color en detalle y si tienen alguna relación con el precio. Ya que como hemos visto en otros estudios analizados en el estado del arte, se han identificado relaciones entre ciertos colores y el precio de venta de las obras de ciertos artistas, por ejemplo, Hosny concluyó que las obras que contiene gris, blanco o negro como color dominante tienden a venderse por cantidades mayores comparadas con otras en las que predominan colores más saturados.

- Otra área de investigación de gran interés para poder contrastar con este estudio sería la aplicación de redes neuronales para la resolución del modelo planteado en este proyecto. Pese a la falta de transparencia que caracteriza a los modelos basados en redes neuronales, también presentan numerosas ventajas como su mejor adaptación “outliers”. Por ello, sería muy interesante poder comparar los modelos resueltos mediante ambas tecnologías y analizar los puntos que mantienen en común o difieren.
- Otra de las áreas que resulta interesante para la realización de una investigación adicional es la relativa a la relación identificada entre la variable que mide la suavidad presente en una obra y su precio de venta. Sería interesante continuar en la investigación de esta relación, utilizando una base de datos con un mayor número de observaciones y variedad con el objetivo de confirmar si esta relación se mantiene en otras áreas como obras de determinados autores.
- Pese a no encontrar una relación clara entre el tamaño de la obra y su precio de venta en el presente estudio, creemos que es interesante seguir investigando esta relación ya que puede existir una relación a igualdad de condiciones. Como podemos comprobar en el siguiente gráfico en el que se muestra la relación entre ambas variables manteniendo el resto de las variables constantes, en el que se puede apreciar cómo puede existir un máximo en el precio para un tamaño cercano al DIN A1 y para tamaños superiores el precio de venta disminuye. No obstante, no se puede concluir que exista una relación ya que el número de observaciones con las que desarrollamos esta conclusión no es suficientemente significativo.

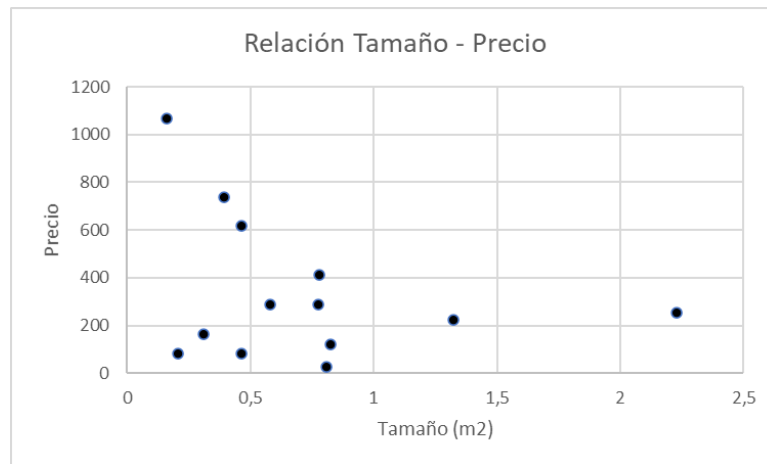


Ilustración 38. Fuente: Elaboración Propia (2021)

6. BIBLIOGRAFÍA

Abidoye, R. B., & Chan, A. P. (Febrero de 2018). *Improving property valuation accuracy: A comparison of hedonic pricing model and artificial neural network*. Pacific Rim Property Research Journal.

Arte, Capital del (10 de Enero de 2018). *Arte figurativo – Características*.

Aubry, M., Kräussl, R., Manso, G., & Spaenjers, C. (Julio de 2020). *Biased Auctioneers*. HEC Paris.

Bellver, J. A., & Martínez, F. G. (2007). *Modelos de valoración económica de obras de arte pictóricas*. Dialnet.

Blasco, I. M. (18 de Febrero de 1998). *Modelos econométricos aplicados a la valoración de bienes inmuebles rústicos*. Dialnet.

Burke, J., & Burns, E. (Marzo de 2021). *What is a neural network? Explanation and examples*. TechTarget.

Fernando, J., & Khartit, K. (13 de Marzo de 2021). *Discounted Cash Flow (DCF)*. Investopedia.

Guillén, A. M. (2015). *Modelo de estimación del precio de las inversiones en obras de arte a través de redes neuronales: La pintura de Picasso en el mercado de las subastas*. Universidad Rey Juan Carlos.

Ghosh, Sourabh, & Mehta, Kiran (2017). *Art as an Investment: Factors Affecting Valuation of Miniature Paintings in India*. Department of Management Studies, IIT Roorkee.

- Hodgson, D., & Galbraith, J. (Junio de 2018). *Econometric Fine Art Valuation by Combining Hedonic and Repeat-Sales Information*. Research Gate.
- Hosny, A. (9 de Diciembre de 2017). *Machine Learning For Art Valuation. An Interview With Ahmed Hosny*. Artnome.
- Koehrsen, W. (28 de Enero de 2018). *Overtting vs. Undertting: A Complete Example*. Patuakhali Science and Technology University.
- Michelone, & López, M. (9 de Julio de 2019). *¿Cómo funcionan las redes neuronales?*. Unocero.
- Rodríguez, L. (2021). *Qué son las redes neuronales y cuál es su aplicación en el marketing*. Artyco.
- Tabales, J. M., Carmona, F. J., & Ocerin, J. M. (29 de Marzo de 2017). *Neural networks (NN) applied to the commercial properties valuation*. Informes de la Construcción.
- Trigo, A (2021). *Criterios para Valorar una Obra de Arte*.
- Zorita, C. M., Álvarez, C., Borrás, F., & Budría, S. (2019). *Modelos Cuantitativos para la Economía y la Empresa en 101 ejemplos*. EV Services.
- Carollo, M. Carmen (2011-2012). *Contrastes de Hipótesis*. Universidad de Santiago de Compostela
- MC Larrosa, Juan (2021). *Econometría: Regresión Hedónica*.

7.ANEXOS

En primer lugar, se presenta la base de datos empleada para el análisis del modelo, el cálculo de los parámetros y el desarrollo del proyecto.

Painting, Ame	351	19	1	1	0	1	0	1	1	1
Painting, Ame	144	19	1	1	0	1	0	1	1	1
Painting, Ame	288	21	1	1	0	1	0	1	1	0
Painting, Ame	493	21	1	1	0	0	1	1	2	0
Painting, Ame	411	20	1	1	0	1	0	1	1	0
Painting, Ame	226	20	1	1	0	0	0	1	3	0
Painting, Dutc	1151	17	1	1	1	1	0	1	2	1
Miniature pai	103	18	1	1	1	1	0	1	2	0
Painting, Ame	288	21	1	1	0	1	0	1	2	0
Painting, Ame	123	20	1	1	1	1	0	1	3	1
Oil painting, W	550	20	1	1	1	1	1	1	3	0
Painting, Sout	164	21	1	1	1	1	0	1	1	1
Painting, Wil	2466	18	3	1	0	0	1	1	2	0
Oil painting, A	25	21	1	1	1	1	0	1	2	1
Oil Painting Re	82	20	1	1	1	1	0	1	1	0
Painting, Mexi	658	19	1	1	0	0	0	1	1	0
Painting, Lee	616	21	3	1	1	1	0	1	4	0
Painting, Ame	255	21	1	1	0	0	0	1	3	0
2415: Thomas	1315	18	1	1	0	0	1	1	2	0
Oil/canvas, Ar	2055	18	1	1	1	1	0	1	1	1
Oil, Polia Pillin	1315	18	1	1	1	0	1	1	3	0
Oil, Louis Heir	1068	19	1	1	0	1	0	1	2	0
Oil/canvas, Ar	82	20	1	1	0	1	0	1	1	0
Oil, Polia Pillin	1151	18	1	1	1	0	1	1	3	1
Painting, Nort	5753	15	2	1	0	1	1	1	1	1
Painting, Russ	1644	19	2	1	0	1	1	1	1	0
Painting, East	6986	16	2	1	1	0	1	1	2	1
Painting, Gern	658	19	2	1	0	0	1	1	2	0
2207: Oil, Geo	247	21	2	1	0	0	1	1	3	1
Painting, Fren	267	21	2	1	0	0	1	1	2	0
Painting, Italia	1397	18	2	1	1	1	1	1	1	1
Painting, Abstr	144	20	4	1	1	0	1	1	3	1
Painting, Abstr	740	21	3	1	0	1	0	1	3	0
Painting, Abstr	103	20	4	1	0	0	1	1	3	0
Abstract oil pe	452	19	4	2	0	1	0	1	4	0
Abstract oil pe	263	19	4	2	0	0	1	1	2	0
"We Are All Ex	140	19	4	2	1	0	1	1	3	0
Katarzyna G?r	123	20	2	2	0	0	1	1	3	1
Abstract oil pe	534	19	2	2	0	1	0	1	4	0
Abstract oil pe	312	19	2	2	0	1	0	1	4	0
Hiremio Garcí	214	21	1	2	0	1	0	2	4	1
62 California l	329	20	1	2	0	0	1	2	3	1
France Outsid	244	20	1	2	1	0	1	2	4	0
Flower Power	188	20	1	2	0	0	1	2	3	0
Mary Frank O	2180	19	2	2	1	1	1	2	4	1
Abstract Origi	41	20	1	3	1	1	1	2	3	1
Acrylic Paintin	164	20	1	3	1	0	1	2	3	0
Abstract Origi	41	20	3	3	1	1	1	2	3	0
Abstract Origi	41	20	3	3	1	1	1	2	3	0
Abstract Origi	41	20	3	3	1	0	1	2	3	0
Abstract Mod	658	20	3	3	0	1	0	2	4	0
25: ABSTRACT	411	20	1	3	0	1	1	2	4	0
7077E: Fantas	205	20	1	3	0	1	0	2	4	0
Abstract Origi	345	20	3	3	1	1	0	2	3	0
Abstract Paint	49	20	1	3	0	1	0	2	4	0
Abstract Paint	70	20	1	3	1	1	0	2	3	0
1951 Jackson	419	20	3	4	0	1	1	1	3	0
Original Abstr	69	20	3	4	0	1	0	1	3	0
3n Cafe Lady	945	20	3	4	0	0	0	1	4	0
Contemporan	920	19	3	4	0	0	0	1	3	0
Original Abstr	1345	19	3	4	0	1	0	2	4	0
Abstract New	2345	20	3	4	0	1	1	2	4	0
Modern abstr	3020	18	3	4	1	1	1	1	3	1
Sung-Ja Lee; 2	103	20	3	5	1	0	0	1	3	1
Sung-Ja Lee; 2	164	20	3	5	0	0	0	1	3	0
Sung-Ja Lee; 2	82	20	3	5	0	0	0	1	3	0

20thC. Japane	41	20	3	5	0	1	1	1	4	0
Sung-Ja Lee; 2	123	20	3	5	1	0	1	1	3	0
18th Century	2260	18	3	5	0	0	1	1	2	0
Painting, 3n sc	226	20	3	5	0	1	0	3	1	0
Eglomise Chin	82	20	3	5	0	1	0	1	1	0
Painting, Hora	534	19	2	5	0	1	0	1	2	0
Painting, cour	390	18	2	5	0	1	0	1	1	0
Painting, Euro	66	21	2	5	0	1	0	1	2	0
Oil painting, S	25	21	2	5	0	1	1	1	1	0
Rabbit Water	247	19	1	2	0	1	0	3	4	0
Leo Meiersdo	176	20	1	2	0	1	0	3	4	0
40s US WC Pa	575	20	4	2	0	1	1	3	1	0
C1850 Origina	136	20	4	2	0	1	0	3	1	0
Original Chine	411	20	3	2	0	1	0	3	2	0
1870 Japanes	378	19	3	2	1	0	0	3	4	0
4n Impala Ori	1080	19	4	4	1	1	0	3	2	0
44 US WC Pai	1050	18	4	2	0	0	1	3	1	1
17C Colonial f	800	17	1	2	0	1	0	1	1	1
Italian School	4320	17	2	2	0	1	1	1	2	1
Large Painting	4700	17	2	2	0	1	0	1	1	1
Antique oil pa	1350	17	2	2	0	1	0	1	1	0
17C Bel Painti	1500	17	3	2	0	0	1	1	2	0
Circle of Sir G	696	17	4	2	1	0	0	1	1	1
Antique 1700	4300	17	2	4	1	1	1	1	1	1
17th Century	2347	17	2	4	1	1	0	2	1	0
Jan Dircksz Bc	1068	17	2	4	1	1	1	1	2	1
Jan Dircksz Bc	2260	17	2	4	1	1	0	1	1	1
Painting, Italia	1559	16	2	1	0	1	1	1	1	1
Oil/canvas/bc	1559	16	4	1	0	1	0	1	1	1
Painting, Euro	1477	16	1	1	1	1	1	1	1	1
Painting, Italia	13949	16	2	1	0	1	1	1	1	0
Painting, Karl	1569	18	1	1	0	0	1	1	2	0
	1858	18	1	1	0	0	1	1	2	0
Carl Von Hass	2271	18	2	1	1	0	1	1	2	0
	1652	19	2	1	1	0	1	1	2	0
Carl Von Hass	2065	18	2	1	0	0	0	1	2	0
	1156	19	2	1	0	0	0	1	2	0
Hunt Slonem	1404	20	1	1	0	0	1	1	4	0
	1569	20	1	1	0	0	1	1	4	0
HUNT SLONEI	1074	20	1	1	0	0	1	1	4	0
	1045	20	1	1	0	0	1	1	4	0
Mid Century C	1074	18	3	2	1	0	0	1	2	1
	2642	18	3	2	1	0	0	1	2	1
Landscape Or	54	21	2	2	0	0	1	1	2	0
	54	21	2	2	0	0	0	1	2	0
Keith Harit - U	1177	18	3	2	1	1	0	2	4	0
	1033	18	1	2	1	1	1	2	4	0
Clemens van c	496	19	2	2	0	1	0	1	1	1
	1486	19	2	2	0	1	0	1	1	1
English Schoo	2829	19	2	2	1	1	1	1	1	1
	2065	17	2	2	1	1	1	1	1	1
19TH C. ANIM	1404	19	2	2	0	1	1	1	1	0
	1321	19	2	2	0	1	1	1	1	0
Gustave Cour	2271	16	2	2	0	1	1	1	2	0
	2477	17	2	2	0	1	1	1	2	0
	3097	17	2	2	0	1	1	1	2	0
19th century c	826	18	1	2	0	1	0	1	1	1
	764	19	1	2	0	1	0	1	1	1
	640	19	1	2	0	1	0	1	1	1
Jig Jag Lady P	363	19	3	2	0	0	1	1	4	0
Vintage Origin	41	20	3	4	1	1	0	1	2	0
	83	20	3	4	1	1	0	1	2	0
Kersch signed	17	20	1	4	0	0	0	1	1	0
	17	20	1	4	0	0	0	1	1	0
Painting, Fren	3306	18	2	1	0	0	1	1	2	0
Painting, Briti	3512	18	2	1	1	1	1	1	1	1
Painting, Euro	2686	18	2	1	1	1	1	1	1	1
Painting, Philli	3512	20	4	1	0	0	1	1	1	0
Painting, Abel	2892	19	2	1	0	0	1	1	2	0
Painting, Euro	2686	17	2	1	0	1	1	1	1	1
German KPM	3099	19	2	1	0	1	1	1	1	0
Oil/canvas, Al	2479	16	4	1	0	0	1	1	2	0
Oil painting, H	3099	16	2	1	0	0	1	1	4	0
Vladimir Gors	331	20	1	2	0	1	0	1	2	0
Oil painting of	413	18	4	2	1	1	0	1	1	0
19th century f	380	19	2	2	1	1	0	1	2	1
French sailing	1	21	2	2	1	1	0	3	2	0
"Dancer" Limi	99	20	1	2	0	0	0	1	1	0
lot of 4 landsc	58	20	1	2	0	0	0	1	2	0
Oil Painting by	331	20	1	2	0	0	0	1	4	0
Folk Art Painti	83	20	1	2	0	1	0	1	1	1
Painting Geor	99	20	1	2	0	1	0	1	1	0
Norman R. Ep	298	20	1	2	0	1	0	1	3	0
Carol Pinsky C	314	20	1	2	0	0	0	1	3	0
Contemporan	231	20	1	2	0	1	0	1	4	0
Antique Marit	331	19	1	2	1	1	0	3	2	0
Lorenz Griffit	66	20	1	2	1	1	0	1	2	1

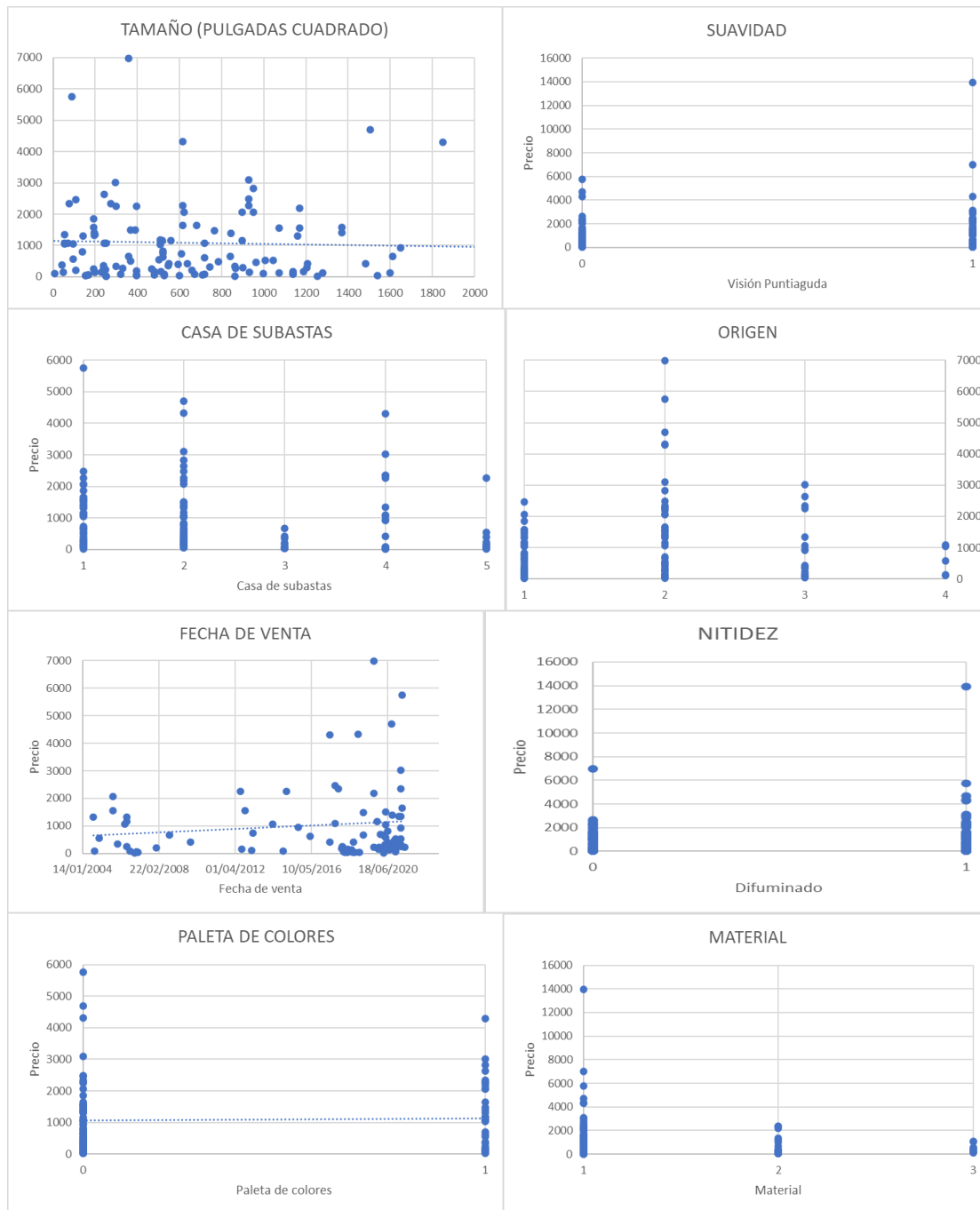
Tabla 7. Fuente: Elaboración Propia (2021)

Para la correcta comprensión de la base de datos, de las gráficas y de las ecuaciones presentadas en este proyecto, se adjunta la codificación empleada para la identificación de las variables:

ORIGEN	AMERICA	1	PALETA DE COLORES	VARIEDAD DE COLORES	0
	EUROPA	2		POCOS COLORES	1
	ASIA	3	NITIDEZ	DIFUMINADO	0
	AFRICA	4		NÍTIDO	1
CASA DE SUBASTAS	CLARS AUCTION GALLERY	1	SUAVIDAD	PUNTIAGUDO	0
	JASPER522	2		REDONDEADO	1
	ARTFORSALE.COM	3	MATERIAL	LIENZO	1
	ROYAL TREASURES WAREHOUSE LLC	4		PAPEL	2
	HUDSON VALLEY AUCTIONS	5		ACUARELA	3
ESTILO	RETRATO/ESCENA	1	CLARIDAD	CLARO	0
	PAISAJE	2			
	ABSTRACTO	3		OSCURO	1
	SURREALISTA	4			

Tabla 8. Fuente: Elaboración Propia (2021)

A continuación, se presentan una serie de gráficas que analizan la base de datos empleada y muestran cómo se distribuye dicha base de datos.



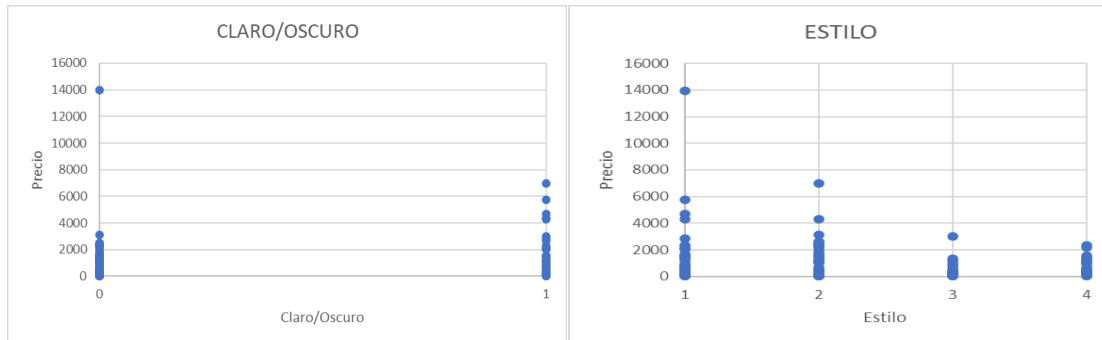


Ilustración 39. Fuente: Elaboración Propia (2021)

De manera complementaria, se adjunta la gráfica de predicción de nuestro modelo para las 153 observaciones que conforman la base de datos. En ella se puede apreciar la predicción utilizando el modelo final, el precio real observado y un intervalo del 95% alrededor del valor de predicción.

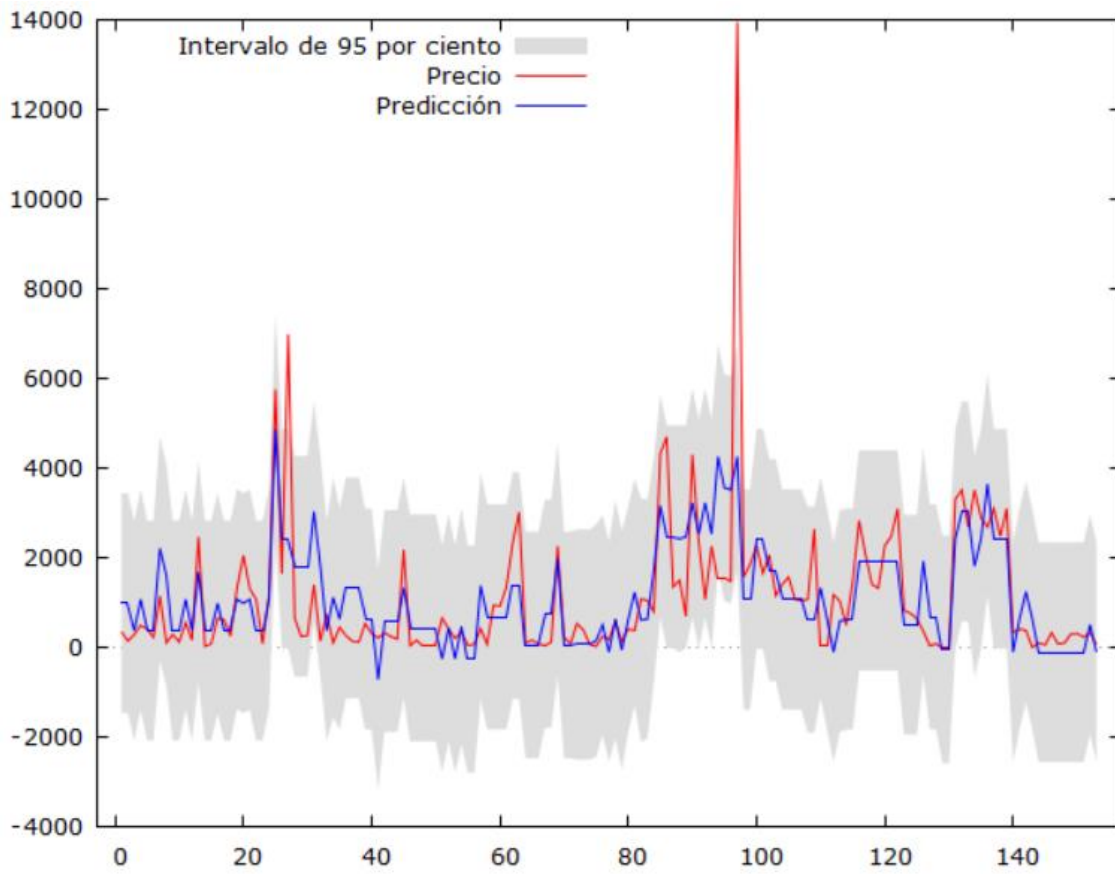


Ilustración 40. Fuente: Elaboración Propia (2021)