



Universidad Nacional de  
Ingeniería

Maestría en Data  
Science:  
Tesis 1: Desarrollo de  
plan de tesis

# Estructura de la revisión de literatura

Mg. Ing. Luis Alberto Sánchez Alvarado

**Se debe seguir los siguientes pasos:**

**1. Búsqueda de  
información**

**2. Organización**

**3. Estructuración**

**4. Escritura**

**Investigación**

**Creación**

# Revisemos la secuencia con un ejemplo



**Título:** Predicción de precio de casas usando machine learning

## **OBJETIVO:**

El objetivo principal de este proyecto de tesis es desarrollar un modelo de machine learning preciso y confiable que pueda predecir el precio de las casas utilizando datos históricos de ventas y características de la propiedad. Se buscará identificar el conjunto óptimo de características predictivas y técnicas de modelado para mejorar la precisión de la predicción del precio de las casas.

## **IMPORTANCIA:**

La predicción precisa del precio de las casas es fundamental tanto para los compradores como para los vendedores en el mercado inmobiliario. Para los compradores, tener una estimación precisa del precio de una propiedad les ayuda a tomar decisiones informadas sobre la compra de una casa y a negociar de manera más efectiva con los vendedores. Para los vendedores, conocer el valor realista de su propiedad les permite fijar un precio competitivo que maximice sus ganancias y minimice el tiempo en el mercado.

## 1. Búsqueda de información

Esta etapa implica recopilar una amplia gama de fuentes relevantes relacionadas con el tema de investigación. La búsqueda de información puede incluir la exploración de bases de datos académicas como science direct. Es crucial identificar y seleccionar las fuentes más autorizadas y actualizadas para garantizar la calidad y la validez de la información recopilada.

## 1. Búsqueda de información

### **EXPLORACIÓN DE BASES DE DATOS ACADÉMICAS:**

Es importante identificar las bases de datos académicas más relevantes como ScienceDirect, Google Scholar, entre otras.



ScienceDirect

Google Scholar

## 1. Búsqueda de información

### **IDENTIFICACIÓN DE PALABRAS CLAVES:**

Las palabras clave son términos específicos o frases que resumen el contenido y el enfoque de tu investigación. Al elegir las palabras clave adecuadas, se puede llevar mejor la búsqueda

# Aplicando: Identificación de palabras claves



**Título:** Predicción de precio de casas usando machine learning


Primero se puede intentar probando con cosas muy generales:

- Forecasting
- Machine learning
- House forecasting

# Aplicando: Identificación de palabras claves







**Título:** Predicción de precio de casas usando machine learning


 ScienceDirect



Search for peer-reviewed journal articles and book chapters (incl


Find articles with these terms  In this journal or book title





 ScienceDirect Journals & Books   Luis Alberto Sanchez Al


Find articles with these terms  

 Advanced search

 Download selected articles  Export


☐ Research article  Full text access

<sup>1</sup> A hybrid machine learning framework for forecasting house price  
Expert Systems with Applications, 14 July 2023  
Choujun Zhan, Yonglin Liu, ... Tommy W. S. Chow  
 View PDF Abstract  Extracts  Figures  Export

☐ Research article  Open access

<sup>2</sup> House price forecasting with neural networks

66,136 results

 Set search alert

Refine by:

Years

☐ 2025 (4)

☐ 2024 (2,429)

☐ ...



# Aplicando: Identificación de palabras claves



**Título:** Predicción de precio de casas usando machine learning

<https://gemini.google.com/app>

Gemini

Consulta las últimas novedades en el [Centro de privacidad de las aplicaciones de Gemini](#)

Hola, Luis  
¿En qué puedo ayudarte?

Ideas para nombres de productos



Plan de estudio de idiomas



Historia de la música



Compara dos cosas



Introduce una petición aquí



# Aplicando: Identificación de palabras claves



**Título:** Predicción de precio de casas usando machine learning

Usemos el siguiente prompt:

Considerar que eres un especialista en investigación y estás desarrollando una tesis para la maestría de ciencia de datos, el proyecto tiene como título: “**Predicción de precio de casas usando machine learning**”. El objetivo de la investigación es “**desarrollar un modelo de machine learning preciso y confiable que pueda predecir el precio de las casas utilizando datos históricos de ventas y características de la propiedad. Se buscará identificar el conjunto óptimo de características predictivas y técnicas de modelado para mejorar la precisión de la predicción del precio de las casas**”. Vas a comenzar con la revisión de artículos en Science Direct , por ello por favor construye fórmulas de búsqueda para encontrar artículos científicos que ayuden con la redacción del estado del arte de la tesis, hazlo en ingles (y al costado pon su equivalencia en español) y usa operadores booleanos.

# Aplicando: Identificación de palabras claves



**Título:** Predicción de precio de casas usando machine learning

Gemini

Consulta las últimas novedades en el [Centro de privacidad de las aplicaciones de Gemini](#)

Mostrar versiones



## Fórmulas de búsqueda para ScienceDirect:

1. Encontrar artículos sobre predicción del precio de viviendas usando machine learning:

- **Fórmula:** "house price prediction" AND (machine learning OR artificial intelligence OR data mining)
- **Equivalente en español:** "predicción del precio de viviendas" AND (machine learning OR inteligencia artificial OR minería de datos)

2. Encontrar artículos sobre el conjunto óptimo de características predictivas para la predicción del precio de viviendas:

- **Fórmula:** "house price prediction" AND "feature selection"
- **Equivalente en español:** "predicción del precio de viviendas" AND "selección de características"

# Aplicando: Identificación de palabras claves



**Título:** Predicción de  
precio de casas usando  
machine learning

1. Encontrar artículos sobre predicción del precio de viviendas usando machine learning:

Fórmula: "house price prediction" AND (machine learning OR artificial intelligence OR data mining)

**Equivalente en español:** "predicción del precio de viviendas" AND (machine learning OR inteligencia artificial OR minería de datos)

2. Encontrar artículos sobre el conjunto óptimo de características predictivas para la predicción del precio de viviendas:

Fórmula: "house price prediction" AND "feature selection"

**Equivalente en español:** "predicción del precio de viviendas" AND "selección de características"

3. Encontrar artículos sobre técnicas de modelado para la predicción del precio de viviendas:

Fórmula: "house price prediction" AND (regression OR classification OR tree OR ensemble)

**Equivalente en español:** "predicción del precio de viviendas" AND (regresión OR clasificación OR árbol OR conjunto)

4. Encontrar artículos sobre el rendimiento de los modelos de machine learning para la predicción del precio de viviendas:

Fórmula: "house price prediction" AND (performance OR accuracy OR error)

**Equivalente en español:** "predicción del precio de viviendas" AND (rendimiento OR precisión OR error)

# Aplicando: Identificación de palabras claves



**Título:** Predicción de precio de casas usando machine learning



ScienceDirect

Journals & Books



Luis Alberto Sanchez A

Find articles with these terms

"house price prediction" AND (machine learning OR artificial intelligence)



Advanced search

69 results

Set search alert

Refine by:

Years

☐ 2024 (9)

☐ 2023 (19)

☐ 2022 (15)

Show more

☐ Download selected articles [Export](#)

☐ Research article Open access

<sup>1</sup> **House Price Prediction using Random Forest Machine Learning Technique**

Procedia Computer Science, 3 February 2022

Abigail Bola Adetunji, Oluwatobi Noah Akande, ... Gbenle Oluwadara

☐ View PDF [Abstract](#) [Export](#)

☐ Research article Open access

<sup>2</sup> **Locally interpretable tree boosting: An application to house price prediction**

Decision Support Systems, 15 October 2023

Anders Hjort, Ida Scheel, ... Johan Pensar

## 1. Búsqueda de información

### **REVISIÓN DE LITERATURA EXISTENTE:**

La revisión de literatura implica identificar y recopilar una amplia gama de fuentes que aborden el tema de investigación específico. Estas fuentes pueden incluir artículos académicos, libros, tesis.

# Aplicando: Revisión de literatura existente



Research article ● Open access

## <sup>1</sup> House Price Prediction using Random Forest Machine Learning Technique

Procedia Computer Science, 3 February 2022

Abigail Bola Adetunji, Oluwatobi Noah Akande, ... Gbenle Oluwadara



[View PDF](#)

[Abstract](#)



[Export](#)



Predicting a price variance rather than a specific value is more realistic and attractive in many real-world applications. Price prediction can be thought of as a classification issue in this situation. However, the House Price Index (HPI) is a common tool for estimating the inconsistencies of house prices. Since housing prices are closely correlated with other factors such as location, city, and population, predicting individual housing prices needs information other than HPI. The HPI is a repeat-sale index that tracks average price shifts in repeat transactions or refinancing of the same assets. Therefore, HPI is ineffective at predicting the price of a single house because it is a rough predictor based on all transactions. This study explores the use of Random Forest machine learning technique for house price prediction. UCI Machine learning repository Boston housing dataset with 506 entries and 14 features were used to evaluate the performance of the proposed prediction model. A comparison of the predicted and actual prices predicted revealed that the model had an acceptable predicted value when compared to the actual values with an error margin of  $\pm 5$ .

# Aplicando: Revisión de literatura existente



Research article ● Open access

## <sup>2</sup> Locally interpretable tree boosting: An application to house price prediction

Decision Support Systems, 15 October 2023

Anders Hjort, Ida Scheel, ... Johan Pensar



[View PDF](#)

[Abstract](#)



[Figures](#)



[Export](#)



We introduce Locally Interpretable Tree Boosting (LitBoost), a tree boosting model tailored to applications where the data comes from several heterogeneous yet known groups with a limited number of observations per group. LitBoost constraints the complexity of a Gradient Boosted Trees model in a way that allows us to express the final model as a set of local Generalized Additive Models, yielding significant interpretability benefits while still maintaining some of the predictive power of a Gradient Boosted Trees model. We use house price prediction as a motivating example and demonstrate the performance of LitBoost on a data set of  $N = 14382$  observations from 15 different city districts in Oslo (Norway). We also test the robustness of LitBoost in an extensive simulation study on a synthetic data set.



# Aplicando: Revisión de literatura existente



Research article ● Open access

<sup>3</sup>

## A comparative assessment of machine learning methods for predicting housing prices using Bayesian optimization

Decision Analytics Journal, 18 January 2023

Salim Lahmiri, Stelios Bekiros, Christos Avdoulas



[View PDF](#)

[Abstract](#) ^

[Figures](#) v

[Export](#) v

The valuation of house prices is drawing noteworthy attention due to worldwide financial and real estate crises in the last decade. Therefore, there is an immediate need to design more effective predictive systems of house prices. Indeed, investors, creditors, and governments are all interested in such predictive systems to improve their buying and lending decisions and activities. This study explores the application of artificial intelligence, machine learning, and nonlinear statistical models to house price prediction problems. In that order, we use boosting ensemble regression trees, support vector regression, and Gaussian process regression. Bayesian optimization is implemented in a ten-fold cross-validation framework to determine their respective optimal kernels and parameter values. Four performance metrics are used to evaluate the prediction ability of each predictive system. The experimental results showed that boosting ensemble regression trees performed the best, followed by Gaussian process regression and support vector regression. In addition, all three aforementioned predictive systems outperformed artificial neural networks and multi-variate regression employed in recent work on the same data set. Under this perspective, it is concluded that boosting ensemble regression trees are clear candidates to be considered for operational house price prediction in Taiwan.

# Aplicando: Revisión de literatura existente

## 4 A hybrid machine learning framework for forecasting house price

Expert Systems with Applications, 14 July 2023

Choujun Zhan, Yonglin Liu, ... Tommy W. S. Chow



[View PDF](#)

[Abstract](#)



[Figures](#)



[Export](#)



House price prediction is one of the most important factors affecting national real estate policies. However, developing an accurate housing price prediction model is a significant challenge for the real estate market. This study presents a framework of house price prediction models that address this issue by improving forecasting performance, explicitly demonstrating the novelty in the Hybrid Bayesian Optimization (HBO) models combined with Stacking (HBOS), Bagging (HBOB), and Transformer (HBOT) techniques. These hybrid models employ Bayesian Optimization for hyperparameter tuning, leading to superior prediction accuracy and stability. Additionally, the proposed framework can assess a statistical and accurate assessment of the predictive performance of house price forecasting models in different scenarios. Furthermore, we constructed a multi-source dataset containing 1,898,175 transactions of the Hong Kong real estate market covering a period from January 2, 1996, to May 13, 2021. This dataset, another major contribution to the field, enables comprehensive model testing and could be a valuable resource for future research. Then, the proposed hybrid models are compared with 18 benchmark models. Thirteen evaluation metrics are used to evaluate the predictive performance, while the non-parametric testing, including Friedman, Iman–Davenport, and Nemenyi post-hoc tests methods, are adopted to assess the significance of differences in the predictive performance of each model. The experimental results show that the HBOS models are superior to the other benchmark models for application in the house price prediction problem. The HBOS-CatBoost model showed superior performance in terms of RMSE compared to both the HBOB-XGBoost and HBOT-ConvLSTM models, with relative RMSE reductions of 5.11% and 25.56%, respectively. The main contributions of this work are the creation of a rich multi-source dataset, the proposal of novel hybrid models for improved house price prediction, and a comprehensive performance evaluation framework. These findings offer a significant step forward in the housing price prediction field.

A large red circle is positioned on the left side of the slide, containing the text '1. Búsqueda de información'.

## 1. Búsqueda de información

### **FILTRADO DE FUENTES:**

Revisar el resumen y ver si es una fuente que podría ser de utilidad para el trabajo de investigación.

# Aplicando: Filtrado de fuentes

## House Price Prediction using Random Forest Machine Learning Technique

### Predicción del precio de la vivienda mediante Random Forest

Predecir una variación de precio en lugar de un valor específico es más realista y atractivo en muchas aplicaciones del mundo real. En esta situación, la predicción de precios puede considerarse como una **cuestión de clasificación**. Sin embargo, el índice de precios de la vivienda (HPI) es una herramienta común para estimar las inconsistencias de los precios de la vivienda. Dado que los precios de la vivienda están estrechamente correlacionados con otros factores como la ubicación, la ciudad y la población, predecir los precios de la vivienda individuales necesita información distinta al HPI. El HPI es un índice de ventas repetidas que rastrea los cambios de precios promedio en transacciones repetidas o refinanciación de los mismos activos. Por lo tanto, el HPI es ineficaz para predecir el precio de una sola casa porque es un predictor aproximado basado en todas las transacciones. Este estudio explora el uso de la técnica de aprendizaje automático Random Forest para la predicción del precio de la vivienda. Se utilizó el conjunto de datos de vivienda de Boston del repositorio de aprendizaje automático de UCI con 506 entradas y 14 características para evaluar el rendimiento del modelo de predicción propuesto. Una comparación de los precios previstos y reales reveló que el modelo tenía un valor previsto aceptable en comparación con los valores reales con un margen de error de  $\pm 5$ .

# Aplicando: Filtrado de fuentes

## Locally interpretable tree boosting: An application to house price prediction

### LitBoost: Una aplicación a la predicción de precios de casas

Presentamos (LitBoost), un modelo de impulso de árboles diseñado para aplicaciones donde los datos provienen de varios grupos heterogéneos pero conocidos con un número limitado de observaciones por grupo. LitBoost limita la complejidad de un modelo de árboles potenciados por gradiente de una manera que nos permite expresar el modelo final como un conjunto de modelos aditivos generalizados locales, lo que produce importantes beneficios de interpretabilidad y al mismo tiempo mantiene parte del poder predictivo de un modelo de árboles potenciados por gradiente. Utilizamos la predicción del precio de la vivienda como ejemplo motivador y demostramos el rendimiento de LitBoost en un conjunto de datos de  $N = 14382$  observaciones de 15 distritos diferentes de la ciudad de Oslo (Noruega). También probamos la solidez de LitBoost en un extenso estudio de simulación en un conjunto de datos sintéticos.

# Aplicando: Filtrado de fuentes

**A comparative assessment of machine learning methods for predicting housing prices using Bayesian optimization**

**Evaluación comparativa de los métodos de machine learning para predecir los precios de la vivienda mediante la optimización bayesiana**

La valoración de los precios de la vivienda está llamando la atención debido a las crisis financieras e inmobiliarias mundiales de la última década. Por tanto, existe una necesidad inmediata de diseñar sistemas predictivos de precios de la vivienda más eficaces. De hecho, los inversores, los acreedores y los gobiernos están interesados en estos sistemas predictivos para mejorar sus decisiones y actividades de compra y préstamo. Este estudio explora la aplicación de inteligencia artificial, aprendizaje automático y modelos estadísticos no lineales a problemas de predicción del precio de la vivienda. En ese orden, utilizamos árboles de regresión de conjuntos potenciadores, regresión de vectores de soporte y regresión de procesos gaussianos. La optimización bayesiana se implementa en un marco de validación cruzada diez veces para determinar sus respectivos núcleos óptimos y valores de parámetros. Se utilizan cuatro métricas de rendimiento para evaluar la capacidad de predicción de cada sistema predictivo. Los resultados experimentales mostraron que la mejora de los árboles de regresión de conjuntos tuvo el mejor rendimiento, seguida de la regresión del proceso gaussiano y la regresión del vector de soporte. Además, los tres sistemas predictivos antes mencionados superaron a las redes neuronales artificiales y la regresión multivariada empleadas en trabajos recientes sobre el mismo conjunto de datos. Bajo esta perspectiva, se concluye que los árboles de regresión conjuntos de impulso son candidatos claros a ser considerados para la predicción operativa del precio de la vivienda en Taiwán.

# Aplicando: Filtrado de fuentes

## A hybrid machine learning framework for forecasting house price

### Un marco híbrido de machine learning para pronosticar el precio de la vivienda

La predicción del precio de la vivienda es uno de los factores más importantes que afectan las políticas inmobiliarias nacionales. Sin embargo, desarrollar un modelo preciso de predicción del precio de la vivienda es un desafío importante para el mercado inmobiliario. Este estudio presenta un marco de modelos de predicción del precio de la vivienda que abordan esta cuestión mejorando el rendimiento de la previsión, demostrando explícitamente la novedad de los modelos de optimización bayesiana híbrida (HBO) combinados con técnicas de apilamiento (HBOS), ensacado (HBOB) y transformador (HBOT). . Estos modelos híbridos emplean optimización bayesiana para el ajuste de hiperparámetros, lo que conduce a una precisión y estabilidad de predicción superiores. Además, el marco propuesto puede evaluar una evaluación estadística y precisa del rendimiento predictivo de los modelos de previsión de precios de la vivienda en diferentes escenarios. Además, construimos un conjunto de datos de múltiples fuentes que contiene 1.898.175 transacciones del mercado inmobiliario de Hong Kong que abarcan un período comprendido entre el 2 de enero de 1996 y el 13 de mayo de 2021. Este conjunto de datos, otra contribución importante al campo, permite pruebas integrales del modelo y podría ser un recurso valioso para futuras investigaciones. Luego, los modelos híbridos propuestos se comparan con 18 modelos de referencia. Se utilizan trece métricas de evaluación para evaluar el rendimiento predictivo, mientras que las pruebas no paramétricas, incluidos los métodos de pruebas post-hoc de Friedman, Iman-Davenport y Nemenyi, se adoptan para evaluar la importancia de las diferencias en el rendimiento predictivo de cada modelo. Los resultados experimentales muestran que los modelos HBOS son superiores a los otros modelos de referencia para su aplicación en el problema de predicción del precio de la vivienda. El modelo HBOS-CatBoost mostró un rendimiento superior en términos de RMSE en comparación con los modelos HBOB-XGBoost y HBOT-ConvLSTM, con reducciones relativas de RMSE del 5,11 % y 25,56 %, respectivamente. Las principales contribuciones de este trabajo son la creación de un rico conjunto de datos de múltiples fuentes, la propuesta de novedosos modelos híbridos para mejorar la predicción del precio de la vivienda y un marco integral de evaluación del desempeño. Estos hallazgos ofrecen un importante paso adelante en el campo de la predicción del precio de la vivienda.

## 2. Organización

Trata de estructurar y ordenar los artículos recopilados según su relevancia y contenido, facilitando así el proceso de análisis y síntesis en etapas posteriores de la revisión del estado del arte.



# Aplicando: organización

1 / 8



100%



Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

**ScienceDirect**

Procedia Computer Science 199 (2022) 806–813

**Procedia**

Computer Science

[www.elsevier.com/locate/procedia](http://www.elsevier.com/locate/procedia)

The 8<sup>th</sup> International Conference on Information Technology and Quantitative Management  
(ITQM 2020 & 2021)

## House Price Prediction using Random Forest Machine Learning Technique

Abigail Bola Adetunji<sup>a</sup>, Oluwatobi Noah Akande<sup>\*b</sup>, Funmilola Alaba Ajala<sup>a</sup>, Ololade Oyewo<sup>a</sup>, Yetunde Faith Akande<sup>c</sup>, Gbenle Oluwadara<sup>b</sup>

<sup>a</sup>Department of Computer Science, Faculty of Computing and Informatics, Ladoke Akintola University of Technology, Nigeria

<sup>b</sup>Computer Science Department, College of Pure and Applied Sciences, Landmark University, Nigeria

<sup>c</sup>Accounting Department, College of Business Sciences, Landmark University, Nigeria

### Abstract

Predicting a price variance rather than a specific value is more realistic and attractive in many real-world applications. Price prediction can be thought of as a classification issue in this situation. However, the House Price Index (HPI) is a common tool for estimating the inconsistencies of house prices. Since housing prices are closely correlated with other factors such as location, city, and population, predicting individual housing prices needs information other than HPI. The HPI is a repeat-sale index that tracks average price shifts in repeat transactions or refinancing of the same assets. Therefore, HPI is ineffective at predicting the price of a single house because it is a rough predictor based on all transactions. This study explores the use

Descargar los archivos y como recomendación se puede nombrar:

Identificación+

Autor+

Año+

Título:

**A1\_Abigail\_2021\_House Price Prediction using Random Forest Machine Learning Technique**

# Aplicando: organización

Decision Support Systems 178 (2024) 114106



Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: [www.elsevier.com/locate/dss](http://www.elsevier.com/locate/dss)



## Locally interpretable tree boosting: An application to house price prediction



Anders Hjort<sup>a,b,\*</sup>, Ida Scheel<sup>a</sup>, Dag Einar Sommervoll<sup>b,c,d</sup>, Johan Pensar<sup>a</sup>

<sup>a</sup> Department of Mathematics, University of Oslo, Norway

<sup>b</sup> Eiendomsverdi AS, Norway

<sup>c</sup> School of Economics and Business, Norwegian University of Life Sciences (NMBU), Norway

<sup>d</sup> NTNU Trondheim Business School, Norway

### ARTICLE INFO

#### Keywords:

Gradient boosted trees

Consolidated addition models

### ABSTRACT

We introduce Locally Interpretable Tree Boosting (LitBoost), a tree boosting model tailored to applications where the data comes from several heterogeneous yet known groups with a limited number of observations per group.

Descargar los archivos y como recomendación se puede nombrar:

Identificación+

Autor+

Año+

Título:

A2\_Hjort\_2024\_Locally interpretable tree boosting An application to house price prediction

# Aplicando: organización

1 / 8



100%



Decision Analytics Journal 6 (2023) 100166



Contents lists available at ScienceDirect

Decision Analytics Journal

journal homepage: [www.elsevier.com/locate/dajour](http://www.elsevier.com/locate/dajour)



## A comparative assessment of machine learning methods for predicting housing prices using Bayesian optimization

Salim Lahmiri <sup>a,\*</sup>, Stelios Bekiros <sup>b,c</sup>, Christos Avdoulas <sup>d,e</sup>

<sup>a</sup> Department of Supply Chain & Business Technology Management, John Molson School of Business, Concordia University, Montreal, Canada

<sup>b</sup> FEMA, University of Malta, MSD 2080 Msida, Malta

<sup>c</sup> LSE Health, London School of Economics and Political Science (LSE), London WC2A2AE, UK

<sup>d</sup> Athens University of Economics and Business, Athens, Greece

<sup>e</sup> Inter-American Research Center, ACHMEA, Canada

### ARTICLE INFO

#### Keywords:

House price prediction  
Predictive analytics  
Boosting ensemble regression trees  
Support vector regression  
Gaussian process regression  
Bayesian optimization

### ABSTRACT

The valuation of house prices is drawing noteworthy attention due to worldwide financial and real estate crises in the last decade. Therefore, there is an immediate need to design more effective predictive systems of house prices. Indeed, investors, creditors, and governments are all interested in such predictive systems to improve their buying and lending decisions and activities. This study explores the application of artificial intelligence, machine learning, and nonlinear statistical models to house price prediction problems. In that order, we use boosting ensemble regression trees, support vector regression, and Gaussian process regression. Bayesian optimization is implemented in a ten-fold cross-validation framework to determine their respective optimal kernels and parameter values. Four performance metrics are used to evaluate the prediction ability of each predictive system. The experimental results showed that boosting ensemble regression trees performed the best, followed by Gaussian process regression and support vector regression. In addition, all three aforementioned

Descargar los archivos y como recomendación se puede nombrar:

Identificación+

Autor+

Año+

Título:

A3\_Lahmiri\_2023\_A comparative assessment of machine learning methods for predicting housing prices using Bayesian optimization

# Aplicando: organización

house price 1 / 15 100%

Expert Systems With Applications 233 (2023) 120981

Contents lists available at ScienceDirect

**Expert Systems With Applications**

journal homepage: [www.elsevier.com/locate/eswa](http://www.elsevier.com/locate/eswa)

**A hybrid machine learning framework for forecasting house price**

Choujun Zhan <sup>a,b</sup>, Yonglin Liu <sup>b</sup>, Zeqiong Wu <sup>b</sup>, Mingbo Zhao <sup>b,c,\*</sup>, Tommy W.S. Chow <sup>c</sup>

<sup>a</sup> School of Computing, South China Normal University, Guangzhou, Guangdong, China  
<sup>b</sup> School of Electrical and Computing Engineering, Nanfang College, Guangzhou, Guangdong, China  
<sup>c</sup> Department of Electronic Engineering, City University of Hong Kong, Hong Kong, China

**ARTICLE INFO**

**Keywords:**  
Machine learning  
House prices  
Stacking  
Bagging  
Transformer

**ABSTRACT**

House price prediction is one of the most important factors affecting national real estate policies. However, developing an accurate housing price prediction model is a significant challenge for the real estate market. This study presents a framework of house price prediction models that address this issue by improving forecasting performance, explicitly demonstrating the novelty in the Hybrid Bayesian Optimization (HBO) models combined with Stacking (HBOS), Bagging (HBOB), and Transformer (HBOT) techniques. These hybrid models employ Bayesian Optimization for hyperparameter tuning, leading to superior prediction accuracy

Descargar los archivos y como recomendación se puede nombrar:

Identificación+

Autor+


Año+

Título:

A4\_Zhan\_2023\_A hybrid machine learning framework for forecasting house price

Otro punto importante es guardar las referencias, que luego será de ayuda para la redacción

# Aplicando: organización

 Research article ● Open access

## <sup>1</sup> House Price Prediction using Random Forest Machine Learning Technique

Procedia Computer Science, 3 February 2022



Abigail Bola Adetunji, Oluwatobi Noah Akande, ... Gbenle Oluwadara

 View PDF   Abstract ▼   Export ▲


### Export

- > Save to RefWorks
- > Export citation to RIS
- > Export citation to BibTeX
- > Export citation to text



 A1\_Abigail\_2021\_House Price Prediction ...  
 ScienceDirect\_citations\_1713901496169

# Aplicando: organización

 Research article ● Open access  
<sup>2</sup> **Locally interpretable tree boosting: An application to house price prediction**

Decision Support Systems, 15 October 2023



Anders Hjort, Ida Scheel, ... Johan Pensar

 View PDF Abstract ▾ Figures ▾ Export ▴

## Export


- > Save to RefWorks
- > Export citation to RIS
- > Export citation to BibTeX
- > Export citation to text



 A2\_Hijort\_2024\_Locally interpretable tree ...  
 ScienceDirect\_citations\_1713901742306

**Supongamos que eliminamos esta fuente:**

# Aplicando: organización

 Research article ● Open access

## <sup>3</sup> A comparative assessment of machine learning methods for predicting housing prices using Bayesian optimization

Decision Analytics Journal, 18 January 2023



Salim Lahmiri, Stelios Bekiros, Christos Avdoulas

 View PDF   Abstract ▾   Figures ▾   Export ▴

### Export

- > Save to RefWorks
- > Export citation to RIS
- > Export citation to BibTeX
- > Export citation to text



 A3\_Lahmiri\_2023\_A comparative assessm...  
 ScienceDirect\_citations\_1713901828177



# Aplicando: organización



Research article ● Full text access

## 4 A hybrid machine learning framework for forecasting house price

Expert Systems with Applications, 14 July 2023

Choujun Zhan, Yonglin Liu, ... Tommy W. S. Chow



View PDF

Abstract ▾

Figures ▾

Export ▴

### Export

> Save to RefWorks

> Export citation to RIS

> Export citation to BibTeX

> Export citation to text



A4\_Zhan\_2023\_A hybrid machine learnin...



ScienceDirect\_citations\_1713901900911

## Aplicando: organización

Finalmente para cerrar esta parte se puede crear un archivo para tomar las primeras notas de los documentos encontrados:

Identificación	Nombre	Año	Autor	Palabra búsqueda	Algoritmos

## 3. Estructuración

Implica organizar los artículos recopilados en secciones coherentes, resumir y analizar críticamente la información, y presentarla de manera sistemática para facilitar la comprensión y la identificación de tendencias, y áreas de oportunidad en la investigación existente.

# Aplicando: Estructuración

Identificación	Nombre	Año	Autor	Palabra búsqueda	Idea principal	Base de datos
A1	A1_Abigail_2021_House Price Prediction using Random Forest Machine Learning Technique	2021	Abigail	"house price prediction" AND (machine learning OR artificial intelligence OR data mining)	El artículo explora la efectividad del algoritmo de aprendizaje automático Random Forest en la predicción de precios de viviendas utilizando el conjunto de datos de viviendas de Boston. Argumenta que predecir un rango de precios en lugar de un valor específico es más práctico e investiga las limitaciones del Índice de Precios de la Vivienda (IPV) para la predicción de precios de viviendas individuales.	datos de vivienda de Boston del repositorio de aprendizaje automático UCI. 506 entradas representa datos agregados sobre 14 características de los hogares. Además se complemento con dat actual
	A3_Lahmiri_2023_A comparative assessment of			"house price prediction" AND (machine learning	Probar alternativas de análisis que den un mejores resultados en la predicción de precios de las casas	Se utilizaron un total de 414 instancias de datos, obtenidas del Ministerio del Interior de Taiwán, con cinco atributos de entrada son la edad de la casa en años, la distancia a la estación de transporte más



## 4. Escritura

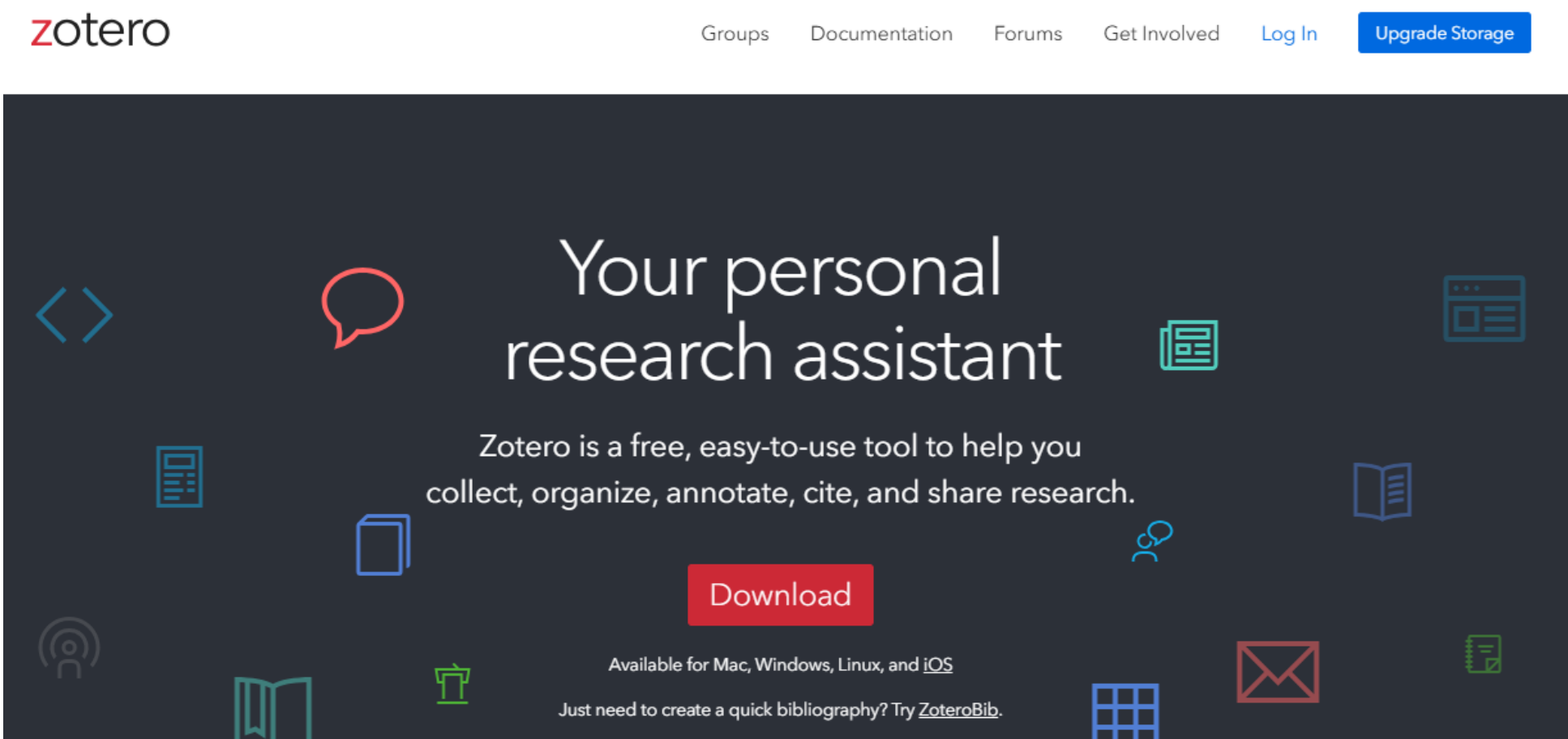
- Integrativo (critica la literatura para introducir nuevas perspectivas, es la más común)
- Temático (tema principal o recurrente)
- Cronológico (interesa la línea de tiempo)
- Metodológico (compara y contrasta diferentes métodos)
  - Teórico (compara y contrasta teorías y modelos)

Para la escritura se necesitará referenciar y para ello se debe usar **ZOTERO**:



# Creación: Escritura (Zotero)

<https://www.zotero.org/>



# Creación: Escritura (Zotero)

[Groups](#)[Documentation](#)[Forums](#)[Get Involved](#)[Log In](#)[Register for a free account](#)[Forgot your password?](#)

## Login

Username or Email

Password

☐ Remember Me

Keep me signed in

Login to Zotero



# Creación: Escritura (Zotero)



The screenshot shows the Zotero website's 'Downloads' page. The header features the Zotero logo, a navigation menu with 'Home', 'Groups', 'Documentation', 'Forums', and 'Get Involved', and a search bar. A blue 'Upgrade Storage' button is in the top right. The main content area has a breadcrumb 'Home > Downloads'. Two large cards are displayed: 'Zotero 6 for Windows' and 'Zotero Connector'. The first card shows the Zotero logo, the title 'Zotero 6 for Windows', the subtitle 'Your personal research assistant', a blue 'Download' button, and a link to 'Other platforms'. The second card shows the Chrome logo plus the Zotero logo, the title 'Zotero Connector', the subtitle 'Save to Zotero from your browser', a blue 'Install Chrome Connector' button, and a description: 'Zotero Connectors automatically sense content as you browse the web and allow you to save it'.

**zotero**

[Home](#) [Groups](#) [Documentation](#) [Forums](#) [Get Involved](#)  [Search](#)

[Upgrade Storage](#)

[Home](#) > Downloads



## Zotero 6 for Windows

Your personal research assistant

[Download](#)

[Other platforms](#)



## Zotero Connector

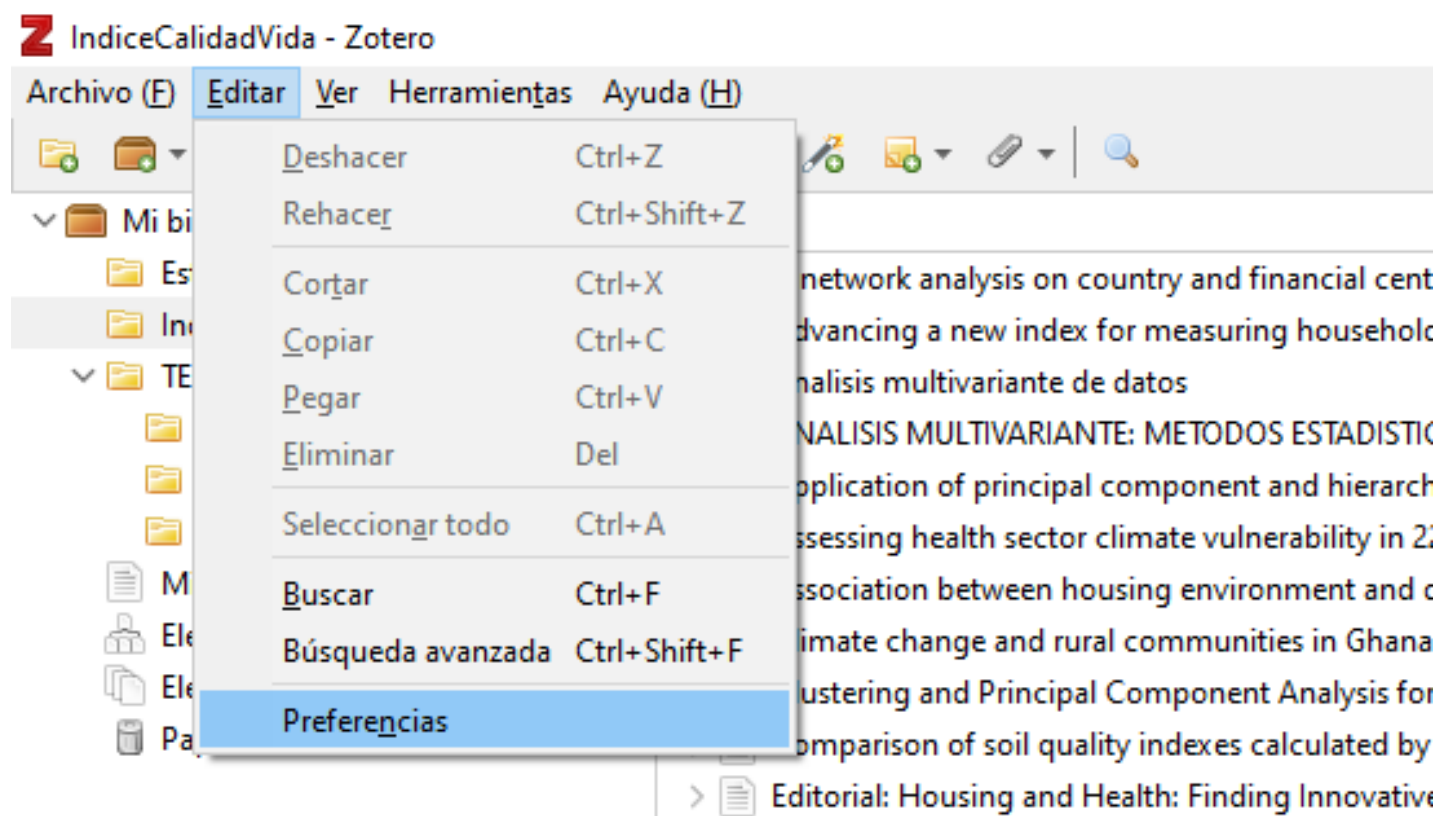
Save to Zotero from your browser

[Install Chrome Connector](#)

Zotero Connectors automatically sense content as you browse the web and allow you to save it

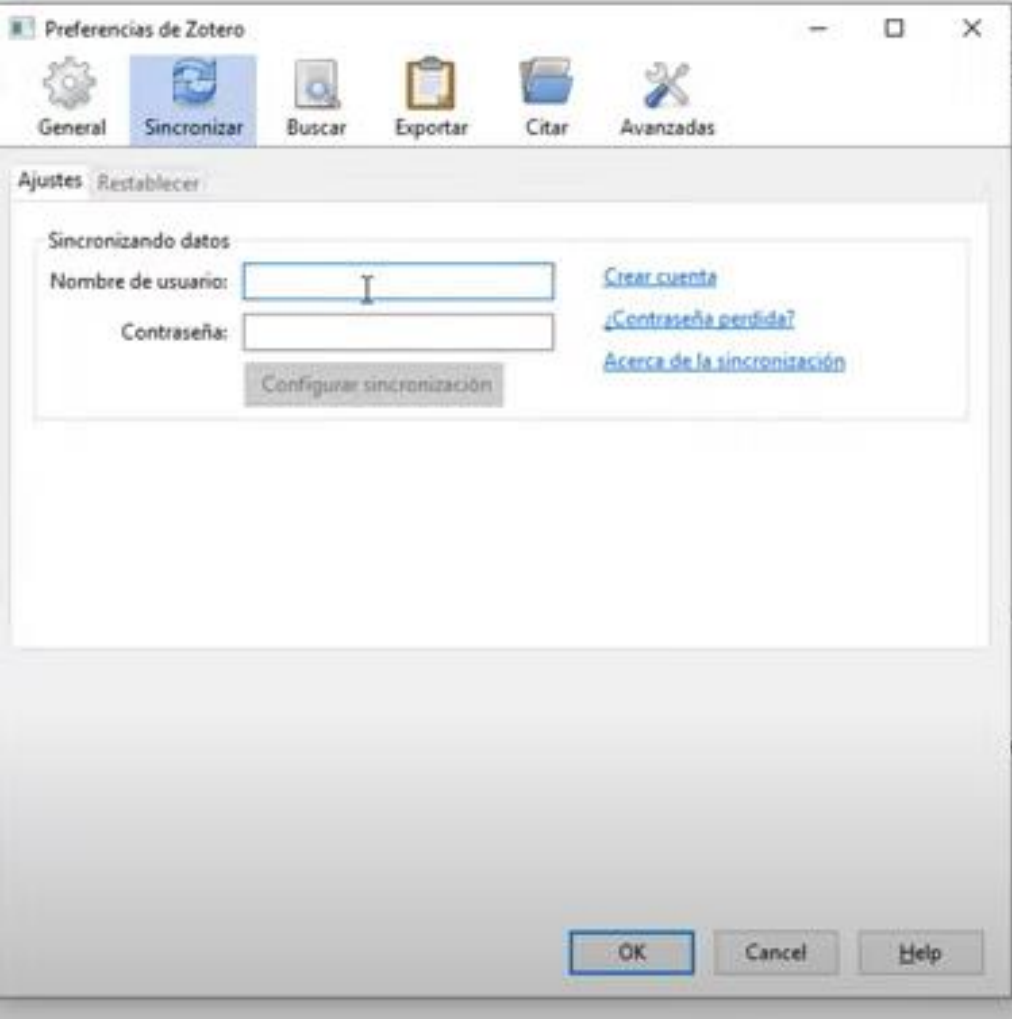
# Creación: Escritura (Zotero)

Una vez instalado ir a la parte de Editar>>Preferencias:



# Creación: Escritura (Zotero)

Sincronizar:



The image shows the 'Preferencias de Zotero' (Zotero Preferences) dialog box, specifically the 'Sincronizar' (Sync) tab. The window has a title bar with standard minimize, maximize, and close buttons. Below the title bar is a toolbar with icons for General, Sincronizar (selected), Buscar, Exportar, Citar, and Avanzadas. The main content area is titled 'Ajustes' (Settings) and includes a 'Restablecer' (Reset) link. Under the 'Sincronizando datos' (Syncing data) section, there are two text input fields: 'Nombre de usuario:' (Username) and 'Contraseña:' (Password). To the right of these fields are three links: 'Crear cuenta' (Create account), '¿Contraseña perdida?' (Forgot password?), and 'Acerca de la sincronización' (About synchronization). Below the password field is a 'Configurar sincronización' (Configure synchronization) button. At the bottom of the dialog are 'OK', 'Cancel', and 'Help' buttons.

Preferencias de Zotero

General Sincronizar Buscar Exportar Citar Avanzadas

Ajustes Restablecer

Sincronizando datos

Nombre de usuario:

Contraseña:

[Crear cuenta](#)

[¿Contraseña perdida?](#)

[Acerca de la sincronización](#)

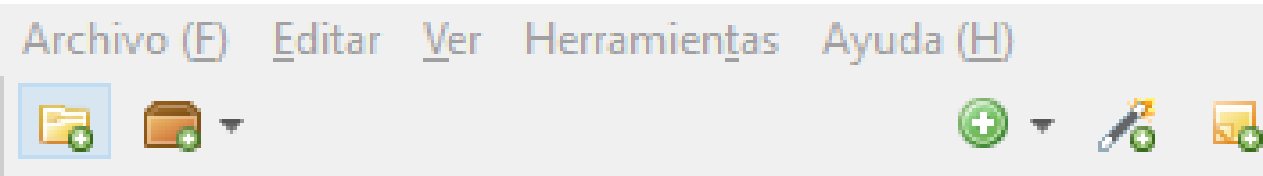
Configurar sincronización

OK Cancel Help

# Creación: Escritura

Ingresar una nueva colección:

 ÍndiceCalidadVida - Zotero



Insertar

Nueva colección



Introduzca un nombre para esta colección:

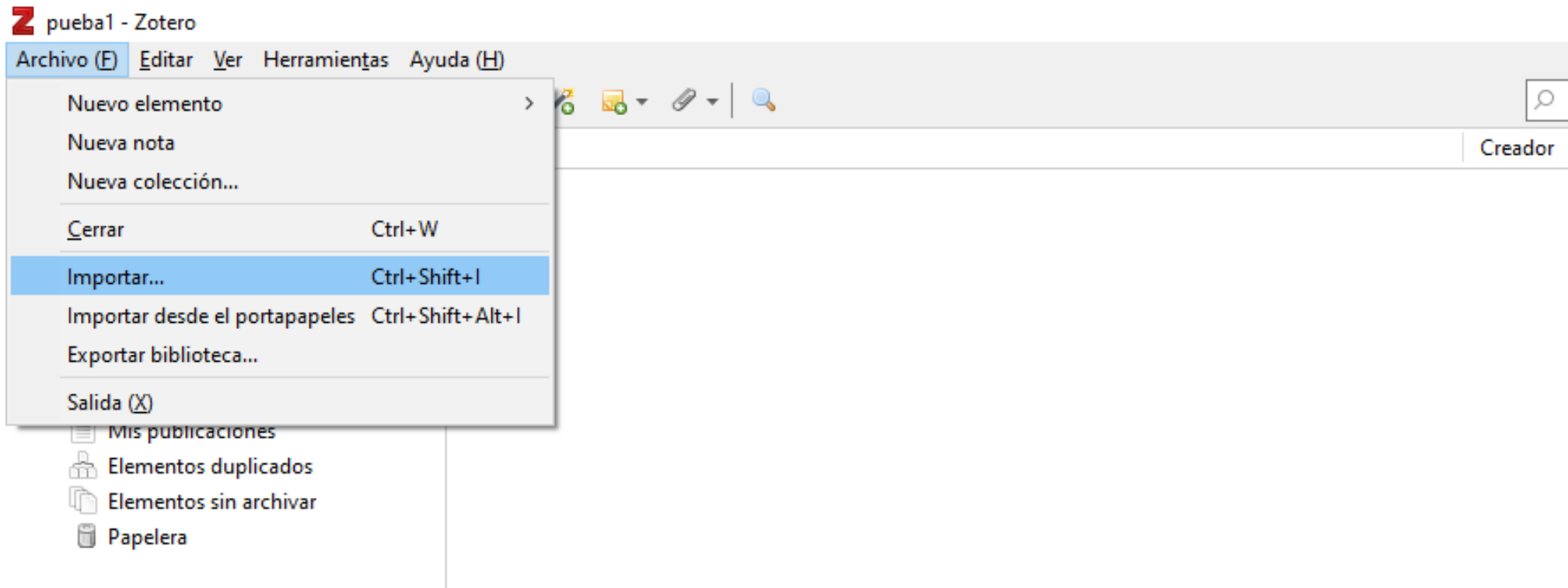
Sin título

OK

Cancel

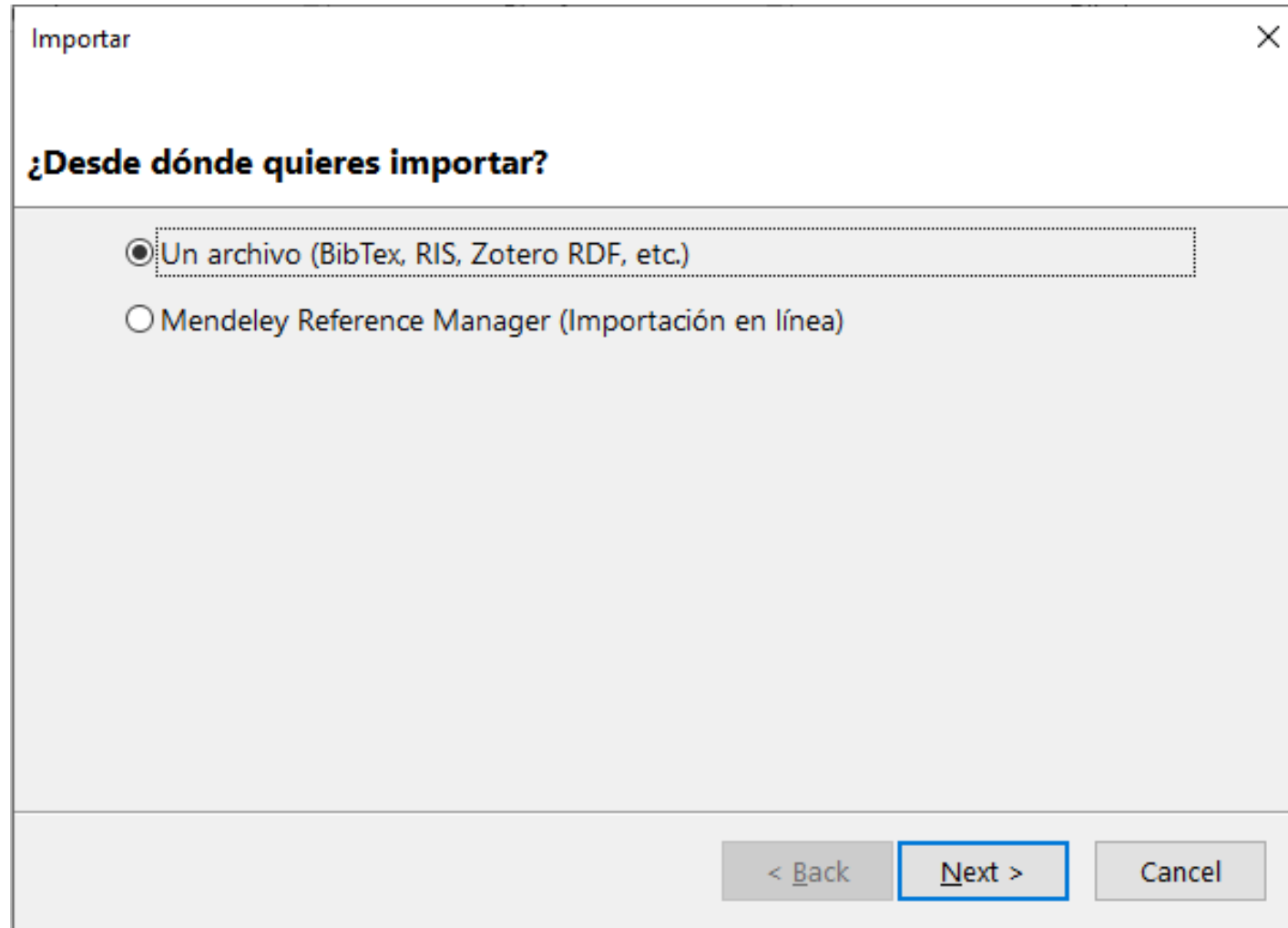
# Creación: Escritura (Zotero)

Y comenzamos incorporando las referencias que se tiene:



# Creación: Escritura (Zotero)

Seleccionamos BibTex:



A screenshot of the 'Importar' (Import) dialog box in Zotero. The window has a title bar with 'Importar' and a close button. The main content area is titled '¿Desde dónde quieres importar?' (Where do you want to import from?). It contains two radio button options: 'Un archivo (BibTex, RIS, Zotero RDF, etc.)' (highlighted with a blue dashed border) and 'Mendeley Reference Manager (Importación en línea)'. At the bottom, there are three buttons: '< Back', 'Next >' (highlighted with a blue border), and 'Cancel'.

Importar

**¿Desde dónde quieres importar?**

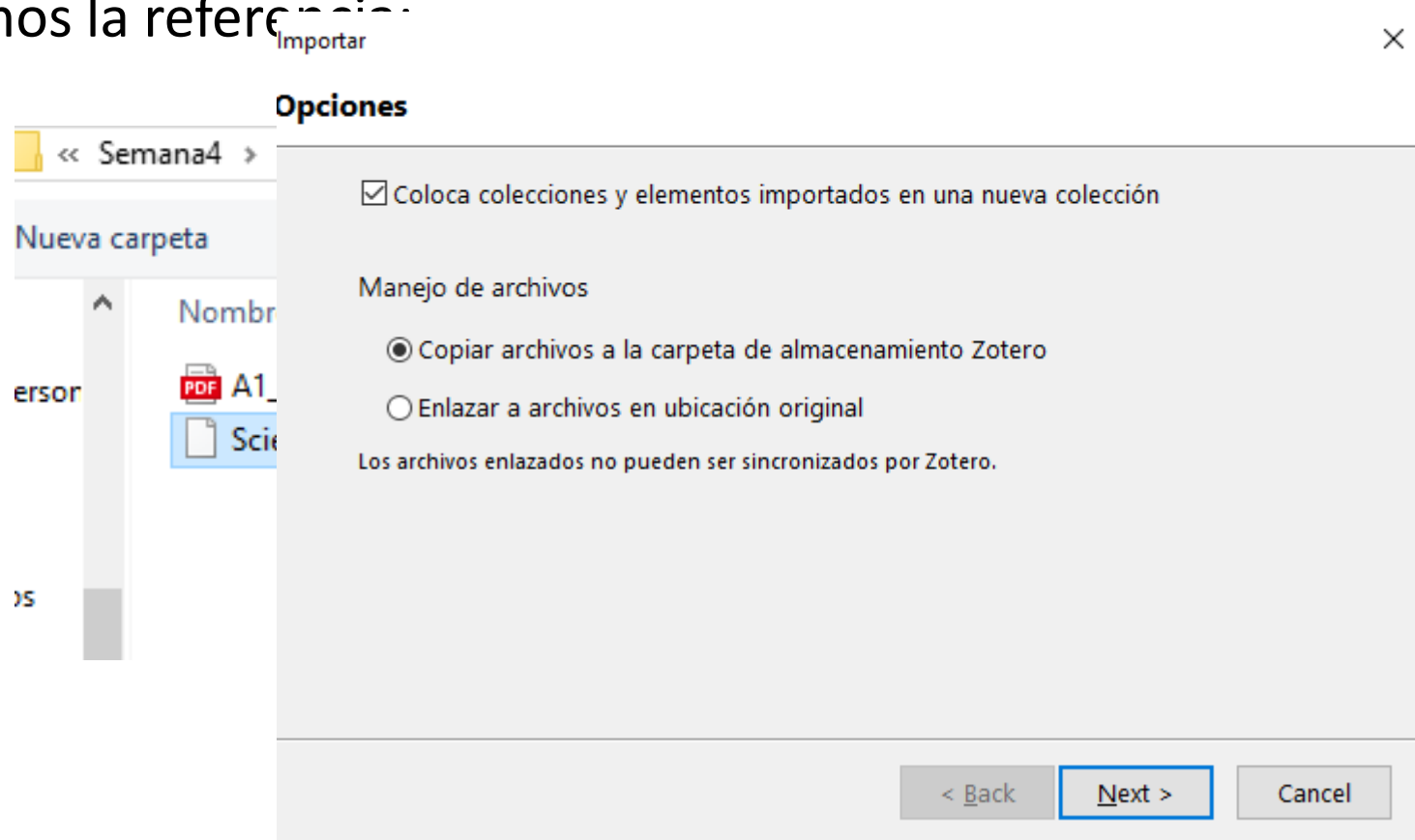
☒ Un archivo (BibTex, RIS, Zotero RDF, etc.)

☐ Mendeley Reference Manager (Importación en línea)

< Back   Next >   Cancel

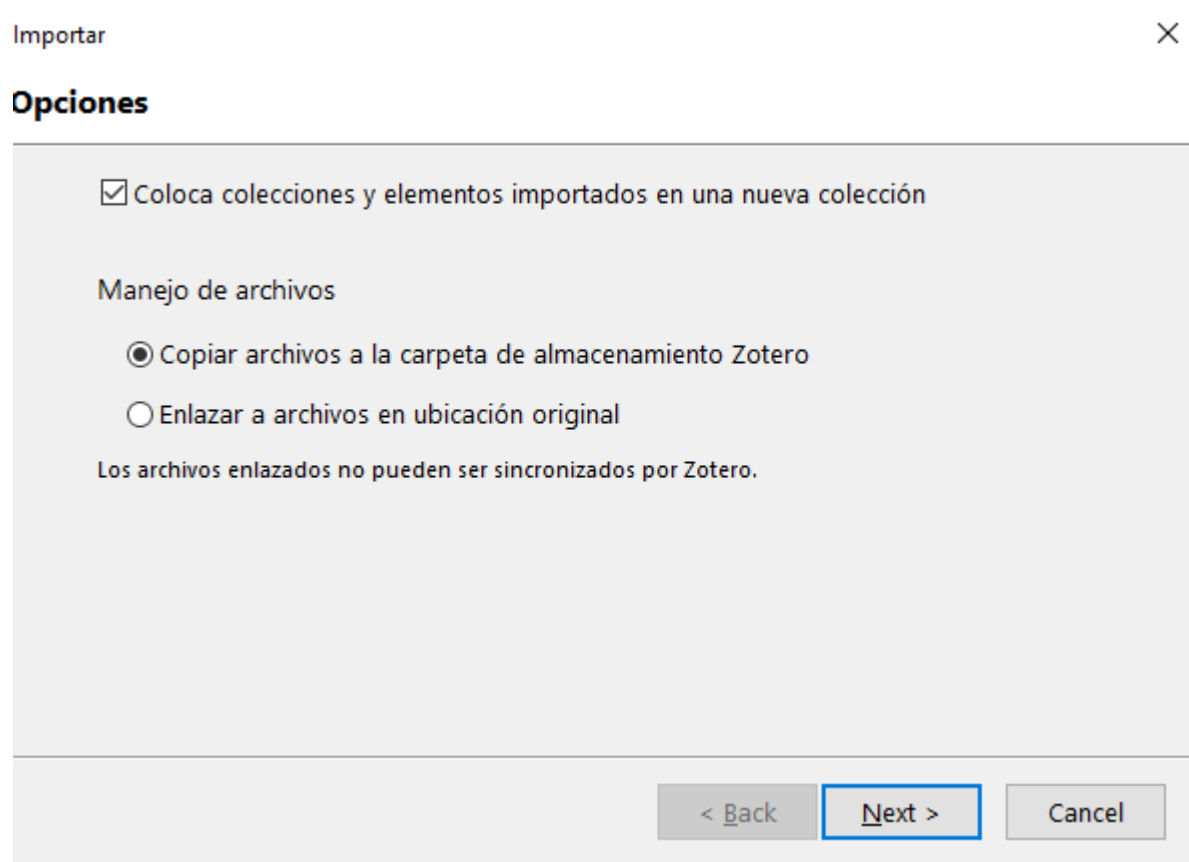
# Creación: Escritura (Zotero)

Seleccionamos la referencia



# Creación: Escritura (Zotero)

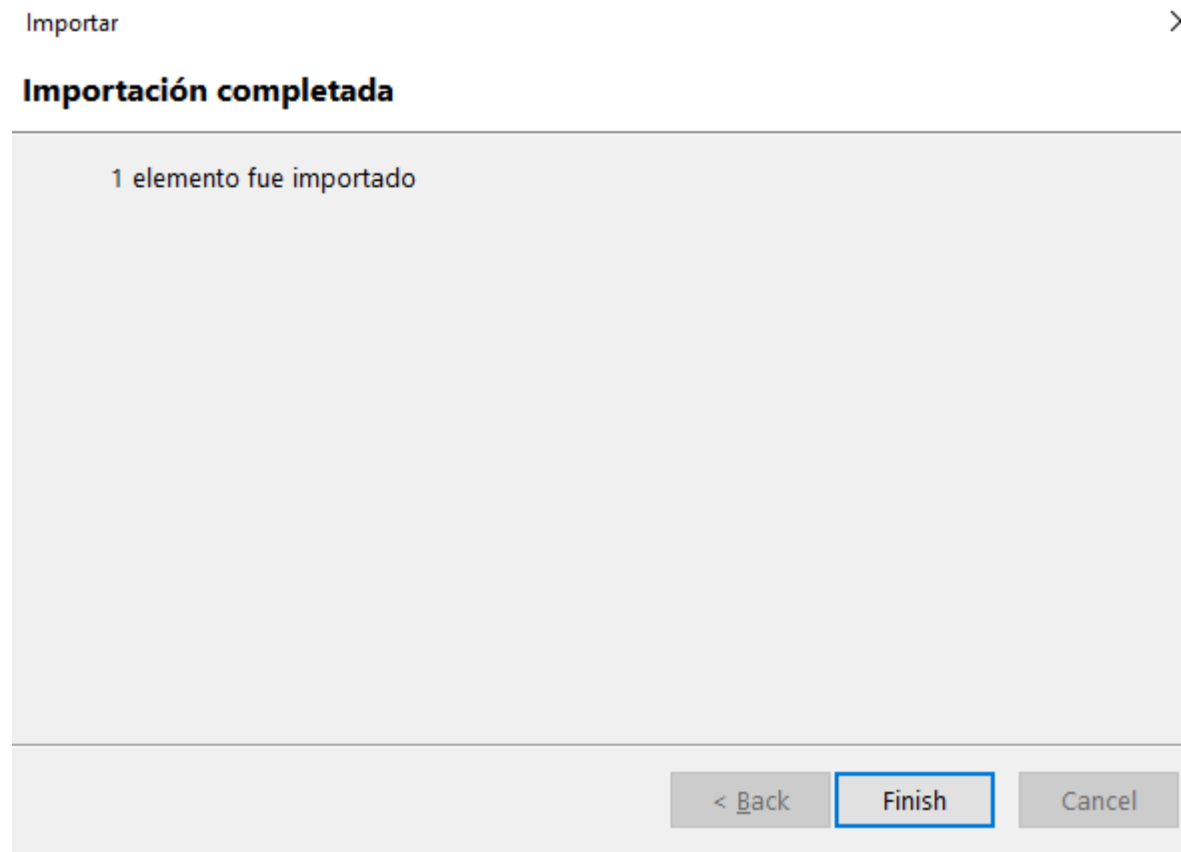
Seleccionamos la referencia:





# Creación: Escritura (Zotero)

Se completa el proceso:



# Creación: Escritura (Zotero)

Seleccionamos la referencia:

The screenshot shows the Zotero 'Create New Item' dialog box. The 'Title' field is filled with 'House Price Prediction using Random Forest Machine Learning Technique' and the 'Creator' field with 'Adetunji et al.'. The 'Information' tab is selected on the right, showing the item type as 'Academic journal article' and a list of authors. The 'Summary' field contains a paragraph of text.

Field	Value
Título	House Price Prediction using Random Forest Machine Learning Technique
Creador	Adetunji et al.
Tipo de elemento	Artículo de revista académica
Título	House Price Prediction using Random Forest Machine Learning Technique
Autor	Adetunji, Abigail ...
Autor	Akande, Oluwato...
Autor	Ajala, Funmilola ...
Autor	Oyewo, Ololade
Autor	Akande, Yetunde ...
Autor	Oluwadara, Gbenle
Resumen	Predicting a price variance rather than a specific value is more realistic and attractive in many real-world applications. Price prediction can be thought of as a classification issue in this situation. However, the House Price Index (HPI) is a common tool for estimating the inconsistencies of house prices. Since housing prices are closely correlated with other factors such as location, city, and population,

# Creación: Escritura (Zotero)

Nueva pestaña x E A comparative assessment of machine learning methods for predicting housing prices using Bayesian optimization x +

← → ↻ [sciencedirect.com/science/article/pii/S277266223000061](https://www.sciencedirect.com/science/article/pii/S277266223000061)

ScienceDirect® Journals & Books Search... Luis Alberto S...

View PDF Download full issue

Outline

- Highlights
- Abstract
- Keywords
- 1. Introduction
- 2. Methods
- 3. Data and results
- 4. Discussion

**Decision Analytics Journal**  
Volume 6, March 2023, 100166

## A comparative assessment of machine learning methods for predicting housing prices using Bayesian optimization

**Extensions**

Acceso completo  
Estas extensiones pueden ver y cambiar información en este sitio web.

- Zotero Connector
- Gestionar extensiones

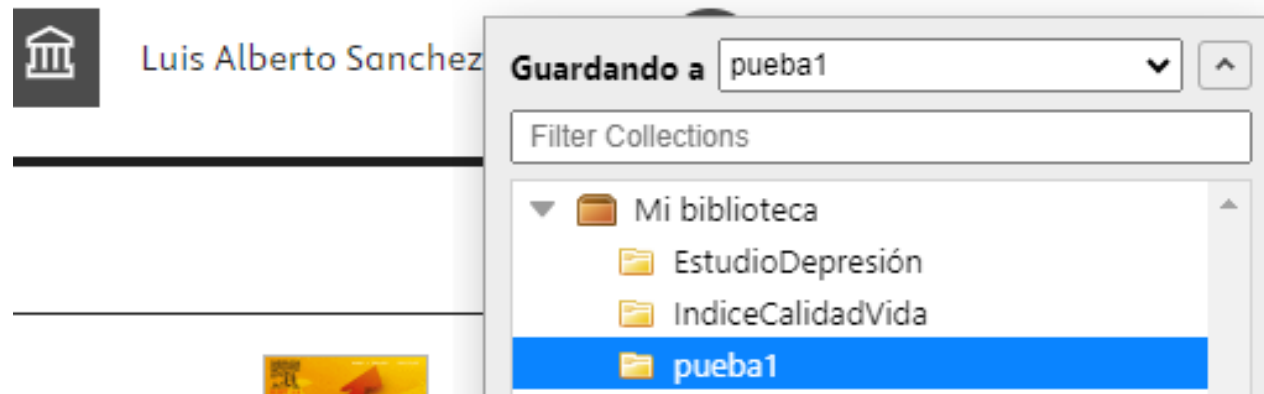
Zotero Connector  
Tiene acceso a este sitio web

**An operational toolbox for solving conflicting claims problems**  
Decision Analytics Journal, Volume 6, 2023, Article...  
Miguel Ángel Mirás Calvo, ..., Estela Sánchez Rodríguez  
 View PDF

**The value of a south-facing orientation: A hedonic pricing analysis of the...**

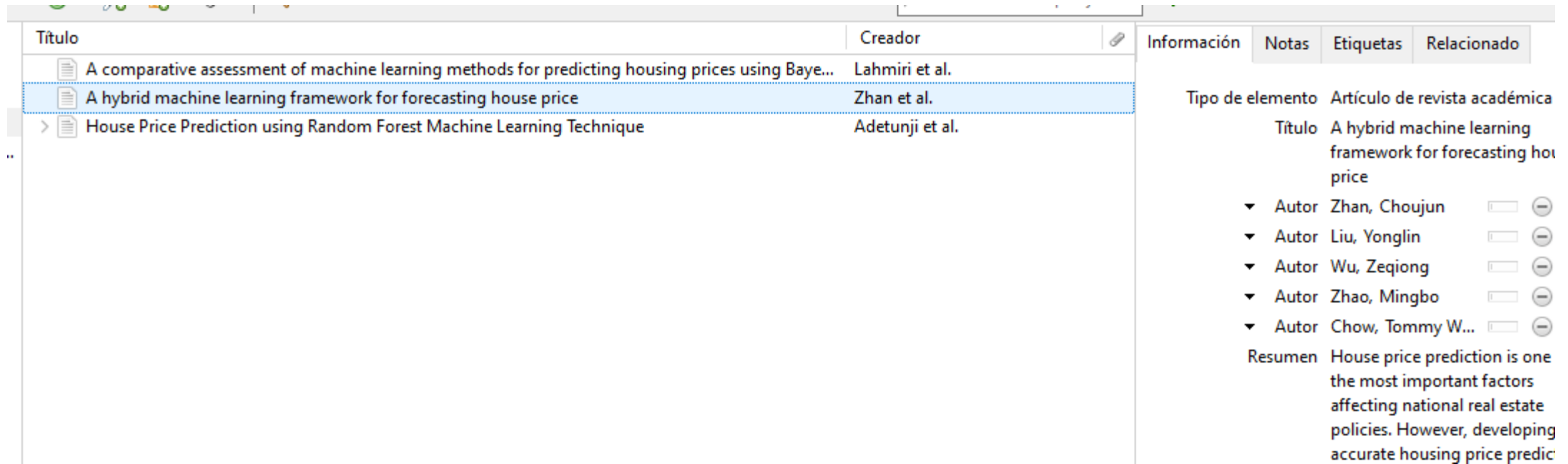
# Creación: Escritura (Zotero)

Se selecciona la carpeta de su proyecto:



# Creación: Escritura (Zotero)

Y con ello se irá completando:



The screenshot shows the 'New Item' dialog box in Zotero. On the left, there is a table with two columns: 'Título' and 'Creador'. The second row is selected, showing the title 'A hybrid machine learning framework for forecasting house price' and the creator 'Zhan et al.'. On the right, there are tabs for 'Información', 'Notas', 'Etiquetas', and 'Relacionado'. The 'Información' tab is active, displaying the following details:

Tipo de elemento	
Tipo de elemento	Artículo de revista académica
Título	A hybrid machine learning framework for forecasting house price
Autor	Zhan, Choujun
Autor	Liu, Yonglin
Autor	Wu, Zeqiong
Autor	Zhao, Mingbo
Autor	Chow, Tommy W...
Resumen	House price prediction is one of the most important factors affecting national real estate policies. However, developing accurate housing price prediction

# Aplicando: Escritura

Ahora lo que resta es hacer la escritura:

2021	Abigail	El artículo explora la efectividad del algoritmo de aprendizaje automático Random Forest en la predicción de precios de viviendas utilizando el conjunto de datos de viviendas de Boston. Argumenta que predecir un rango de precios en lugar de un valor específico es más práctico e investiga las limitaciones del Índice de Precios de la Vivienda (IPV) para la predicción de precios de viviendas individuales.
2023	Lahmiri	Probar alternativas de análisis que den un mejores resultados en la predicción de precios de las casas, y que superen a las redes neuronales y análisis multivariado que se venían usando
2023	Zhan	Indica que para la predicción de precios han existido métodos que han trabajado con modelos de regresión, luego con el pasar de lo años se llegó al uso de modelos de machine learning que dieron buenos resultados que los métodos tradicionales, en tiempos más recientes los modelos híbridos ofrecen dar el mejor resultado

El estudio de modelos para predecir los precios de viviendas es un tema comúnmente abordado en la literatura. Inicialmente, los enfoques más tradicionales han sido los métodos de análisis multivariados, como señala Lahmiri (2023). Sin embargo, con la aparición del machine learning, se han obtenido resultados prometedores según lo mencionado por Zhan (2023), quien indica que ahora los modelos híbridos son una opción viable. Estas investigaciones resaltan que el uso de machine learning ha mejorado significativamente la capacidad predictiva del precio de la vivienda. Por ejemplo, Abigail (2024) demuestra en su estudio que incluso es posible generar intervalos de predicción, lo cual amplía las posibilidades de aplicación y precisión en este campo

# Aplicando: Escritura

Ahora lo que resta es hacer la escritura:

2021	Abigail	datos de vivienda de Boston del repositorio de aprendizaje automático UCI. 506 entradas representa datos agregados sobre 14 características de los hogares. Además se complemento con dat actual
2023	Lahmiri	Se utilizaron un total de 414 instancias de datos, obtenidas del Ministerio del Interior de Taiwán, con cinco atributos de entrada son la edad de la casa en años, la distancia a la estación de transporte más cercana en metros, el número de tiendas de conveniencia en el círculo de vida a pie, y las coordenadas
2023	Zhan	Usan una base del mercado inmobiliario de Hong Kong y el sitio web de la Autoridad Monetaria de Hong Kong. Contiene 1,898,175 registros, con atributos de la viviendas y variables económicas

Las investigaciones encontradas en el área de predicción de precios de viviendas utilizan una variedad de variables, tanto aquellas directamente relacionadas con las características de las viviendas como factores externos. Además, estas investigaciones abordan el análisis tanto con conjuntos de datos relativamente pequeños como con datos históricos de gran tamaño, como lo evidencia el estudio de Zhan (2023), el cual emplea un conjunto de datos masivo que contiene 1,898,175 registros

# Aplicando: Escritura

Ahora lo que resta es hacer la escritura:

2021	Abigail	Aplica normalización de los datos y codificación para las variables categóricas
2023	Lahmiri	
2023	Zhan	Usando boxplot identifica valores atípicos, y también hace una recodificación de las variables categóricas, normalización de

En las metodologías revisadas, es común la aplicación de técnicas de normalización de datos y un análisis exploratorio exhaustivo. Este análisis incluye la identificación de valores atípicos utilizando herramientas gráficas como el box plot.



# Aplicando: Escritura

Ahora lo que resta es hacer la escritura:

2021	Abigail	Random Forest	.Se hace una exploración de los datos, incluyendo una matriz de correlaciones. Se corre el modelo y se revisa las métricas. Se valida con un gráfico del valor actual y el predicho	Error Absoluto Medio (MAE) Error Cuadrático Medio (RMSE)
2023	Lahmiri	ERT (Boosting Ensemble Regression Trees) Gaussian Process Regression SVR (Support Vector	Revisión exploratoria de los datos, Se aplicó el algoritmo de optimización bayesiana para encontrar los parámetros óptimos para los modelos de regresión	MAE, RMSE, MARE, MAPE
2023	Zhan	tres tipos diferentes de modelos híbridos (HBOS, HBOB y HBOT)	Se aplicó el algoritmo de optimización bayesiana	Usa 13 métricas, entre ellas MSE, RMSE, MAE, R2, MAD, MAPE

En cuanto a los algoritmos utilizados en la predicción de precios de viviendas, el Random Forest suele destacarse como uno de los más efectivos. Sin embargo, también se han explorado modelos híbridos, entre los cuales el más exitoso ha sido identificado como el HBOST, según Zhan (2023). Estos modelos suelen implementar una optimización de parámetros basada en el método bayesiano. Entre las métricas utilizadas para evaluar su desempeño se destacan el MAE (Error Absoluto Medio) y el RMSE (Error Cuadrático Medio)

## Aplicando: Escritura

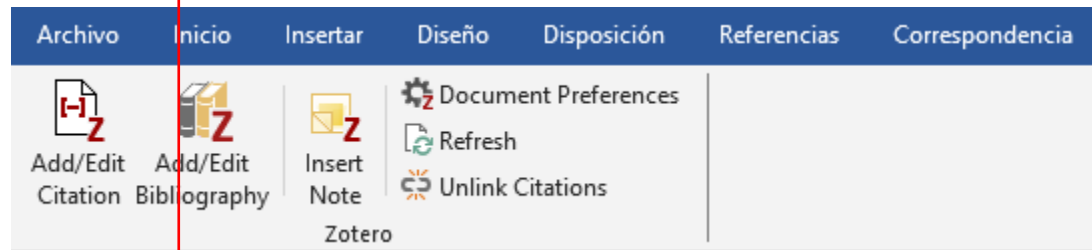
El estudio de modelos para predecir los precios de viviendas es un tema comúnmente abordado en la literatura. Inicialmente, los enfoques más tradicionales han sido los métodos de análisis multivariados, como señala Lahmiri (2023). Sin embargo, con la aparición del machine learning, se han obtenido resultados prometedores según lo mencionado por Zhan (2023), quien indica que ahora los modelos híbridos son una opción viable. Estas investigaciones resaltan que el uso de machine learning ha mejorado significativamente la capacidad predictiva del precio de la vivienda. Por ejemplo, Abigail (2024) demuestra en su estudio que incluso es posible generar intervalos de predicción, lo cual amplía las posibilidades de aplicación y precisión en este campo.

Las investigaciones encontradas en el área de predicción de precios de viviendas utilizan una variedad de variables, tanto aquellas directamente relacionadas con las características de las viviendas como factores externos. Además, estas investigaciones abordan el análisis tanto con conjuntos de datos relativamente pequeños como con datos históricos de gran tamaño, como lo evidencia el estudio de Zhan (2023), el cual emplea un conjunto de datos masivo que contiene 1,898,175 registros.

En las metodologías revisadas, es común la aplicación de técnicas de normalización de datos y un análisis exploratorio exhaustivo. Este análisis incluye la identificación de valores atípicos utilizando herramientas gráficas como el box plot.

En cuanto a los algoritmos utilizados en la predicción de precios de viviendas, el Random Forest suele destacarse como uno de los más efectivos. Sin embargo, también se han explorado modelos híbridos, entre los cuales el más exitoso ha sido identificado como el HBOST, según Zhan (2023). Estos modelos suelen implementar una optimización de parámetros basada en el método bayesiano. Entre las métricas utilizadas para evaluar su desempeño se destacan el MAE (Error Absoluto Medio) y el RMSE (Error Cuadrático Medio).

# Aplicando: Escritura ... Para las citas usar Zotero



## Zotero - Preferencias del documento

### Estilo de cita:

American Chemical Society  
American Medical Association 11th edition  
American Political Science Association  
American Psychological Association (APA) 7th edition  
American Sociological Association (ASA) 6th/7th edition  
Chicago Manual of Style 17th edition (author-date)  
Chicago Manual of Style 17th edition (full note)  
Chicago Manual of Style 17th edition (note)

[Gestionar estilos...](#)

Idioma: Español (España)

### Guardar citas como:

☒ Campos (recomendado)

Los campos no se pueden compartir con LibreOffice.  
El documento debe ser guardado como .doc o .docx.

☐ Marcadores

Los marcadores se pueden compartir entre Word y LibreOffice, pero pueden causar errores si se modifican accidentalmente y no se pueden insertar en notas a pie de página.  
El documento debe ser guardado como .doc o .docx.

☒ Actualizar citas automáticamente

La desactivación de actualizaciones puede acelerar la inserción de citas en documentos grandes. Haga clic en Actualizar para actualizar las citas manualmente.

OK

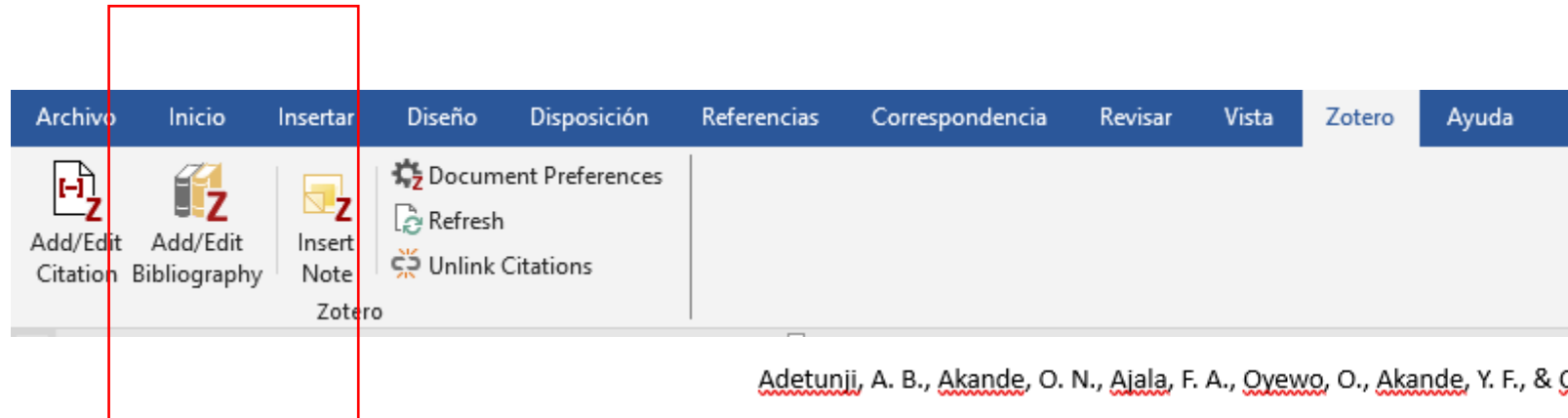
Cancel

Help

# Aplicando: Escritura ... Para las citas usar Zotero

Z zhan				
Mi biblioteca				
<b>A hybrid machine learning framework for forecasting house price</b> Zhan et al. (2023), <i>Expert Systems with Applications</i> , 233, 120981.				
<b>Hierarchical cluster and principal component analyses of multi-scale pore structure...</b> Zhang et al. (2024), <i>Journal of Asian Earth Sciences</i> , 261, 106001.				
<b>Use of machine learning approach to predict depression in the elderly in China: A lo...</b> Su et al. (2021), <i>Journal of Affective Disorders</i> , 282, 289-298.				
Regression SVR (Support				

# Aplicando: Escritura ... Con ello el índice saldrá la bibliografía



Adetunji, A. B., Akande, O. N., Ajala, F. A., Oyewo, O., Akande, Y. F., & Oluwadara, G. (2022).

House Price Prediction using Random Forest Machine Learning Technique. *Procedia*

*Computer Science*, 199, 806-813. <https://doi.org/10.1016/j.procs.2022.01.100>

Lahmiri, S., Bekiros, S., & Avdoulas, C. (2023). A comparative assessment of machine learning

methods for predicting housing prices using Bayesian optimization. *Decision Analytics*

*Journal*, 6, 100166. <https://doi.org/10.1016/j.dajour.2023.100166>

## Otro ejemplo:

La enseñanza en el uso de modelos estadísticos para el estudio de un problema real incluye el uso de gráficos, tablas estadísticas y estadísticos resumen.  
(Angus, T. pp.73-77, 2007).

Angus, T. (2007). *Investigating the Nature of Teacher Knowledge Needed and Used in Teaching Statistics*. Tesis Doctoral en Educación. Massey University.

## Otro ejemplo:

que los alumnos tengan dificultad en reconocer el modo particular de pensamiento estadístico en el cual la modelización juega un papel muy importante.

(Biehler, R. 1999 p.261-262).

Biehler, R. (1999). Discussion: Learning to think statistically and to cope with variation. *International Statistical Review*, 67(3), pp. 259–262.

## Otro ejemplo:

Los estudiantes no están en condiciones de apreciar la relevancia del diseño para el proceso de la generación de datos hasta que no se familiaricen con el análisis de datos. La evidencia mostró que “si tu enseñas diseño antes que análisis de datos, es difícil que los estudiantes entiendan por qué es importante el diseño”.

Cobb, P. et al. (p. 376, 2004)

Cobb, P. & McClain, K. (2004). *Principles of Instructional Design for Supporting the Development of Students' Statistical Reasoning*. Cap 16. En *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*. Kluwer Academic Publisher. Springer.



## Otro ejemplo:

Los estudiantes no están en condiciones de apreciar la relevancia del diseño para el proceso de la generación de datos hasta que no se familiaricen con el análisis de datos. La evidencia mostró que “si tu enseñas diseño antes que análisis de datos, es difícil que los estudiantes entiendan por qué es importante el diseño”.

Cobb, P. et al. (p. 376, 2004)

Cobb, P. & McClain, K. (2004). *Principles of Instructional Design for Supporting the Development of Students' Statistical Reasoning*. Cap 16. En *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*. Kluwer Academic Publisher. Springer.

## Otro ejemplo:

“... cuando se enseña estadística a no estadísticos es importante enfatizar los datos y conceptos a expensa de reducir la teoría y echar mano a problemas de la vida real donde los estudiantes sean parte de la recolección o creación de datos. Usar datos reales y focalizarse en el aprendizaje del estudiante en lugar de la lectura convertirá a los estudiantes en más activos y puede ser visto como aprendizaje centrado en el estudiante”.

Wilberg, M. (2009)

Wilberg, M. (2009). *Teaching Statistics in Integration with Psychology*. Journal of Statistics Education. 17,1.

# Escritura

Es claro que la educación no ha orientado la enseñanza en el uso de modelos estadísticos para el estudio de un problema real. El uso de modelos apropiados incluye gráficos, tablas estadísticas y estadísticos resumen (Angus, T. pp.73-77, 2007). Es poco usual que la enseñanza de la estadística en la universidad haga énfasis en la modelización. Esto puede explicar el hecho de que los alumnos tengan dificultad en reconocer el modo particular de pensamiento estadístico en el cual la modelización juega un papel muy importante (Biehler, R. 1999 p.261-262). Desde otra perspectiva, pero con la misma conclusión Cobb, P. et al. (p. 376, 2004) sustentan que los estudiantes no están en condiciones de apreciar la relevancia del diseño para el proceso de la generación de datos hasta que no se familiaricen con el análisis de datos. La evidencia mostró que “si tu enseñas diseño antes que análisis de datos, es difícil que los estudiantes entiendan por qué es importante el diseño”. Inclusive, los estudiantes del primer año pueden tratarse como no estadísticos en el sentido de Wilberg, M. (2009) al asegurar que cuando se enseña estadística a no estadísticos es importante enfatizar los datos y conceptos a expensa de reducir la teoría y echar mano a problemas de la vida real donde los estudiantes sean parte de la recolección o creación de datos. Usar datos reales y focalizarse en el aprendizaje del estudiante en lugar de la lectura convertirá a los estudiantes en más activos y puede ser visto como aprendizaje centrado en el estudiante. Esto motiva al estudiante lo cual redundará en mejor rendimiento y aprendizaje.

## Segundo trabajo

Mostrar la secuencia de la formación del estado del arte para su proyecto de tesis y redactarlo en el documento de proyecto de tesis.

