

<i>Asignatura</i>	<i>Datos del alumno</i>	<i>Fecha</i>
<i>Módulo 1. Modelos de Inteligencia Artificial</i>	Apellidos: Florindo Lopez	30/12/24
	Nombre: Ivonis	

Proyecto de predicción de Gentrificación Futura en la provincia de Madrid

1. Comprensión del Negocio

Contexto

El aumento de los precios de los alquileres en zonas específicas de Madrid está provocando el desplazamiento de comunidades vulnerables al limitar el acceso a la vivienda. Este fenómeno, conocido como gentrificación, afecta a la cohesión social y la inclusión. Además de generar otros perjuicios, incluso mentales, según la OMS.

Problema

En Madrid, la falta de herramientas predictivas dificulta la aplicación de medidas de forma temprana por instituciones y autoridades locales como el Observatorio de la Exclusión Social en la Comunidad de Madrid (OEISM) o la Fundación FOESSA.

Objetivos

Este proyecto se basa en desarrollar un modelo de IA que pueda predecir el riesgo de Gentrificación, por zonas, en la provincia de Madrid. Para ello el modelo deberá:

- » Identificar patrones de gentrificación mediante el análisis de tendencias históricas
- » Predecir las zonas donde aumentarán los precios del alquiler
- » Facilitar las decisiones informadas de las instituciones para prevenir la exclusión social.

Justificación de la IA

El uso de la IA en ciencias sociales se plantea como un reto debido a la necesidad de garantizar la ética y respetar la nueva regulación europea. No obstante, los beneficios que proporciona son enormes.

- » Puede identificar de patrones complejos más eficientemente.
- » Analiza múltiples variables y se anticipa a los cambios de forma escalable
- » Puede mejorar con nuevos datos
- » Minimiza sesgos y errores humanos propiciando decisiones informadas

Criterios de éxito

Cuantitativos

- » Predicción de zonas con aumentos significativos de precios en las simulaciones con una precisión del 80%
- » Evaluar en retrospectiva: precisión del 85% en 5 años.
- » Predecir con un 70% de precisión las zonas en las que los precios suban un 10%

Cualitativos:

- » Proveer de información útil a instituciones
- » Mejorar la comprensión del fenómeno a nivel académico
- » Buena adopción de la plantilla mediante un dashboard interactivo para personal no técnico...

2. Solución de IA

Modelo de IA

En el presente proyecto se plantea el desarrollo de un modelo de IA basado en algoritmos de aprendizaje supervisado, como Random Forest y regresión lineal, a partir del Dataset de Idealista y entrenado con las variables clave `price`, `district`, `subdistrict`, `floor_area` y `last_update`. Se propone añadir datasets externos de organismos oficiales para enriquecer los análisis. Por último, y también de gran importancia, se pretende que el sistema se despliegue mediante un servidor y un frontend intuitivo y asequible. En la tabla 1 se describen detalladamente los componentes del sistema planteado.

Componentes del Sistema IA

Tabla 1. Caracterización de los componentes del modelo presentado.

Componente	Descripción	Función que realiza	Herramientas Técnicas
Modelo de IA	Algoritmo predictivo basado en aprendizaje supervisado, Random Forest y regresión lineal.	Identificar patrones y zonas con aumentos significativos.	Python (scikit-learn, pandas, numpy), TensorFlow (en caso de necesitar complejizar el entrenamiento).
Dataset principal	Datos de alquiler de Idealista con variables clave (price, district, subdistrict, flor_area, last_update).	Base para entrenar el modelo, analizar tendencias y patrones relacionados con gentrificación.	Idealista, pandas para manipulación de datos.
Fuentes externas	Datasets públicos y oficiales (pendiente estudio): INE, Open Data Madrid, etc., sobre demografía, infraestructuras, renta per cápita y crecimiento económico.	Enriquecer el análisis con contexto social y económico.	Python (pandas para manipulación de datos), APIs de acceso a datos, bases en CSV o JSON.
Interfaz de usuario (UI)	Dashboard interactivo diseñado para personal no técnico, con mapas de calor y reportes visuales para facilitar la comprensión de las predicciones.	Facilita decisiones informadas por instituciones sociales y autoridades locales al ser accesible a personal no técnico.	Streamlit, Google Maps API, Power BI (web), Tableau (web), plotly.Desarrollo de filtros y botones claros, popups de ayuda, etc.
Backend	Sistema que gestiona los cálculos previos del modelo de IA y conecta los datos procesados con la interfaz de usuario para que sean accesibles y utilizables.	Soporta el trabajo técnico previo y asegura que las predicciones estén disponibles para la toma de decisiones en la UI.	AWS, Google Cloud, o Microsoft Azure; Flask/Django para servidores.

3. Evaluación de la situación

Estado del Arte

Informes del FOESSA destacan una carencia de herramientas predictivas en ciencias sociales. De hecho, en el estudio “Artificial Intelligence for Urban Transformation” del MIT Technology Review, se confirma que la IA en gentrificación está en etapas iniciales, centrada en estudios descriptivos y retrospectivos.

Para el entrenamiento del modelo será necesario explorar otros datasets para verificar que cumplen con la normativa y analizar si las variables (densidad poblacional, accesibilidad, crecimiento económico, etc.) contribuyen a alcanzar los objetivos del proyecto.

Validación Ética y Legal

Según el Reglamento (UE) 2024/1689 estamos ante un proyecto de IA que no es de riesgo inaceptable ya que el dataset contiene solo datos agregados y anonimizados, de acceso público. El modelo, además, será entrenado mediante un procedimiento transparente y supervisado que garantice la no discriminación.

Respecto a la ética, esta propuesta ayuda a prevenir la exclusión social mediante predicciones basadas en datos objetivos, evitando sesgos, para mejorar las decisiones orientadas a políticas inclusivas. Además, está orientada al personal no técnico, lo cual lo convierte en un sistema que favorece la equidad y la transparencia.

Plan de Proyecto

Tabla 2. Desglose de gastos económicos para garantizar la funcionalidad técnica, accesibilidad y mantenimiento del sistema presentado en el proyecto.

Concepto	Detalle	Coste Estimado (€)
Desarrollo de Modelo de IA	Entrenamiento del modelo basado en Random Forest y regresión lineal; incluye optimización de hiperparámetros y validación cruzada.	10,000
Adquisición de datos	Acceso a datasets externos como Idealista, INE y Open Data Madrid; preparación y limpieza de datos.	3,000
Infraestructura técnica	Implementación en servidores en la nube (AWS, Azure) para procesar y almacenar datos de manera escalable.	2,500/año
Desarrollo del Dashboard	Creación de una interfaz interactiva con mapas de calor y reportes accesibles para personal no técnico.	5,000
Backend	Configuración del sistema para conectar el modelo de IA con la interfaz de usuario; gestión de consultas y predicciones en tiempo real.	5,000
Mantenimiento anual	Actualización de datasets, ajustes en el modelo y soporte técnico recurrente.	4,000/año

4. Conclusiones

En este proyecto se ha desarrollado una solución IA con componentes éticos y técnicos para implementar un modelo que permita predecir el riesgo de gentrificación en la provincia de Madrid. El diseño se basa en los principios de transparencia y accesibilidad, priorizando una experiencia amigable para el personal no técnico.

Se ha podido notar que, siguiendo la metodología CRISP-DM, algunos elementos del proyecto, como variables nuevas, son indispensables para alcanzar los objetivos. En tanto la viabilidad del proyecto no esté consolidada, se requerirán iteraciones futuras.

Estas iteraciones incluirán la evaluación de la calidad y relevancia de los datasets externos, escogiendo aquellos cuyas variables den robustez a los análisis socioeconómicos del modelo. Además, se incluirán métodos claros de medición de resultados para que las instituciones locales puedan justificar el impacto del proyecto.

En última instancia, el proyecto refuerza su compromiso con la prevención de la exclusión social y el apoyo a instituciones locales mediante una herramienta robusta, accesible y transparente. Asimismo, contribuye al avance en ciencias sociales mediante predicciones a medio y largo plazo.