



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

XAI 3 INTERPRETABLE MODELS

Model Agnostic Methods

21-05-2024

Damián López Blasco
Adrià Comes Sanchis

1. Introducción

1.1 Contexto

Este informe aborda el uso de métodos explicativos independientes del modelo (XAI: *Model-agnostic methods*) para analizar y visualizar la influencia de varias características sobre las predicciones realizadas por modelos de machine learning. En particular, se enfoca en la técnica de gráficos de dependencia parcial (**PDP**), una herramienta que permite entender el impacto marginal de una o más variables independientes en el resultado predicho por un modelo.

El documento se estructura en tres ejercicios prácticos, cada uno diseñado para aplicar **PDP** en diferentes contextos y con diferentes datasets:

- **Unidimensional Partial Dependency Plot:** En este ejercicio, se utiliza un modelo de random forest para predecir el número de alquileres de bicicletas (*cnt*) basado en datos de un conjunto de características. La tarea implica generar un **PDP** unidimensional para visualizar cómo variables como los días desde 2011, la temperatura, la humedad y la velocidad del viento afectan las predicciones del modelo.
- **Bidimensional Partial Dependency Plot:** Este ejercicio expande el análisis al generar un **PDP** bidimensional, combinando la humedad y la temperatura para predecir el número de alquileres de bicicletas. La visualización incluye la distribución de densidad de ambas características en un gráfico 2D, utilizando técnicas como *geom_tile()* para evitar huecos en el gráfico.
- **PDP para explicar el precio de una casa:** Aplicando los conceptos previamente discutidos, este ejercicio se enfoca en predecir el precio de una casa utilizando un modelo de random forest basado en características como el número de habitaciones, baños, el área habitable en pies cuadrados (*sqft_living*), el tamaño del lote (*sqft_lot*), el número de pisos y el año de construcción (*yr_built*). Se requiere generar PDPs para analizar la influencia de estas características en el precio predicho.

Cada ejercicio incluye una serie de preguntas que guían el análisis e interpretación de los resultados obtenidos mediante **PDP**, promoviendo una comprensión profunda de las relaciones entre las características y las predicciones del modelo. Además, se enfatiza la importancia de utilizar herramientas de control de versiones (*git*) y el respaldo del trabajo en plataformas como *GitHub*, asegurando la trazabilidad y colaboración en el desarrollo del análisis.

1.2 Objetivo del informe

El objetivo de este informe es aplicar y demostrar la utilidad de métodos explicativos independientes del modelo (*XAI: Model-agnostic methods*), específicamente mediante el uso de gráficos de dependencia parcial (**PDP**), para interpretar y visualizar la influencia de diversas características en las predicciones realizadas por modelos de machine learning. A través de una serie de ejercicios prácticos, el informe busca:

- **Explicar la variabilidad en predicciones de modelos complejos:** Mostrar cómo los **PDP** pueden descomponer y visualizar la contribución de variables individuales y conjuntas en la predicción de resultados en diferentes contextos.
- **Facilitar la interpretación de modelos de machine learning:** Proveer herramientas y metodologías que permitan a los analistas y científicos de datos entender mejor las relaciones aprendidas por los modelos, promoviendo decisiones más informadas y transparentes.
- **Desarrollar habilidades prácticas en el uso de PDP:** Guiar a los usuarios en la implementación de **PDPs** unidimensionales y bidimensionales, utilizando ejemplos concretos como la predicción de alquileres de bicicletas y precios de viviendas.
- **Promover buenas prácticas en la gestión de proyectos de datos:** Integrar el uso de sistemas de control de versiones (*git*) y plataformas de respaldo (*GitHub*) para asegurar la reproducibilidad y colaboración en los proyectos de análisis de datos.

En resumen, este informe pretende capacitar a los usuarios en la aplicación de técnicas de interpretación de modelos, mejorando la comprensión y confianza en los modelos de machine learning a través de visualizaciones claras y análisis detallados.

2. Metodología

En la sección de metodología de este informe, hemos seguido un enfoque sistemático para nuestro análisis. El código utilizado para realizar la evaluación fue desarrollado en R y está cuidadosamente documentado en un archivo R Markdown. Este archivo se incluye junto con el informe como parte del paquete de entrega. Está disponible para su revisión, lo que garantiza la transparencia y la reproducibilidad de los resultados presentados. Al proporcionar acceso al código, invitamos a la revisión crítica y fomentamos la verificación de las metodologías aplicadas, reforzando así la integridad y la credibilidad de nuestros hallazgos.

3. Análisis

3.1 Aplicación de gráficos de dependencia parcial

En este apartado, exploraremos el uso de gráficos de dependencia parcial (**PDP**) unidimensionales para interpretar los resultados de un modelo de regresión enfocado en la predicción del número de alquileres de bicicletas. Utilizaremos un modelo de *random forest* para esta tarea y generaremos PDPs para visualizar el efecto marginal de diferentes características en el resultado predicho.

Específicamente, analizaremos cómo los días transcurridos desde 2011, la temperatura, la humedad y la velocidad del viento influyen en el número de bicicletas alquiladas. A través de esta visualización, podremos entender mejor las relaciones que el modelo ha aprendido y cómo estas variables individuales afectan las predicciones de alquiler de bicicletas.

[Ver Gráfico unidimensional de dependencia parcial](#)

En el análisis de las condiciones climáticas y su influencia en el número de bicicletas alquiladas, se observan las siguientes tendencias:

Temperatura: Se observa que el número de bicicletas alquiladas aumenta con la temperatura hasta aproximadamente 20 grados Celsius. Sin embargo, cuando la temperatura supera los 25 grados Celsius, el número de alquileres comienza a disminuir.

Cabe recalcar que las líneas son más gruesas de 10 a 25 grados lo que indica que hay más densidad de población y por tanto debemos tomar más importancia a esa franja pues al haber más muestras es más representativo.

Humedad: En relación con la humedad, el número de bicicletas alquiladas se mantiene constante hasta un 50% de humedad. A partir de este punto, a medida que la humedad aumenta, el número de alquileres disminuye.

Por la naturaleza de los datos hay más densidad de población cuando la humedad está entre 40 y 80 por tanto hay que coger con pinzas las muestras que están fuera de este rango pues podrían ser datos atípicos.

Velocidad del Viento: La velocidad del viento también tiene una influencia significativa en las predicciones de alquiler de bicicletas. El número de alquileres disminuye gradualmente hasta alcanzar una velocidad de aproximadamente 23 km/h, a partir de la cual las predicciones se estabilizan.

En cambio aquí la densidad si toma un papel relevante pues nos ayuda a corroborar que a mayor viento menos bicis alquiladas.

Días Transcurridos desde 2011: En términos de la variable temporal, existe una tendencia general al alza en los alquileres de bicicletas con el paso del tiempo. No obstante, las predicciones más recientes indican una disminución en el número de alquileres.

Aquí la densidad de población está igualmente distribuida y por tanto las conclusiones extraídas representan la realidad.

Estas observaciones proporcionan una comprensión detallada de cómo diferentes condiciones climáticas y el paso del tiempo afectan el alquiler de bicicletas, lo que puede ser útil para optimizar la gestión de recursos y mejorar las estrategias de servicio.

3.2 Gráficos de dependencia parcial bidimensional

A continuación nos enfocaremos en la generación de un Partial Dependency Plot (PDP) bidimensional, con el objetivo de comprender cómo la combinación de la humedad y la temperatura influye en el número de bicicletas alquiladas. Utilizaremos un modelo de random forest y, debido al tamaño de los datos, extraeremos un conjunto de muestras aleatorias de la base de datos antes de generar el PDP. La visualización incluirá la distribución de densidad de ambas variables de entrada utilizando un gráfico 2D, siguiendo las técnicas mostradas en las diapositivas de clase y utilizando `geom_tile()` para evitar huecos en el gráfico.

[Ver gráfico de dependencia parcial bidimensional](#)

Después de observar el gráfico podemos observar las siguientes observaciones clave:

Máximo Alquiler de Bicicletas: El mayor número de bicicletas alquiladas se observa cuando la temperatura está entre 15 y 20 grados Celsius y la humedad relativa se encuentra entre 0% y 70%. Este rango de condiciones climáticas parece ser el más favorable para el alquiler de bicicletas.

Efecto de la Humedad Alta y Temperatura Baja: Se nota una disminución en el número de bicicletas alquiladas a medida que la humedad aumenta por encima del 70% y la temperatura desciende por debajo de los 15 grados Celsius. Estas condiciones parecen ser menos atractivas para los usuarios de bicicletas.

Zona de Alquiler Moderado: Entre los 20 y 25 grados Celsius, con una humedad relativa entre 50% y 75%, se observa una zona con un alquiler moderado de bicicletas, indicando que estas condiciones son aceptables pero no las más óptimas.

Tendencias Generales: En general, el número de bicicletas alquiladas tiende a disminuir tanto con un aumento significativo en la humedad como con temperaturas muy bajas o muy altas.

Este gráfico permite visualizar claramente cómo la interacción entre la temperatura y la humedad afecta las predicciones del modelo respecto al alquiler de bicicletas, proporcionando información valiosa para entender y optimizar el uso de bicicletas en función de las condiciones climáticas.

3.3 PDP para explicar el precio de un hogar

En este apartado, aplicaremos los conceptos previamente discutidos para predecir el precio de una casa utilizando el dataset *kc_house_data.csv*. Emplearemos un modelo de *random forest* para realizar la predicción basada en las características: número de dormitorios, número de baños, superficie habitable (*sqft_living*), tamaño del lote (*sqft_lot*), número de pisos y año de construcción (*yr_built*).

Para visualizar las relaciones que el modelo ha aprendido, generaremos gráficos de dependencia parcial (**PDP**). Dado el tamaño del dataset, extraeremos un conjunto de muestras aleatorias antes de generar los PDPs, asegurando así la eficiencia del proceso.

El objetivo es analizar la influencia de las características mencionadas (dormitorios, baños, superficie habitable y número de pisos) en el precio predicho de las casas, proporcionando una comprensión más profunda de cómo cada variable contribuye a las predicciones del modelo.

[Ver gráfico de dependencia parcial \(precio hogar\)](#)

En el análisis de las características estructurales de las casas y su influencia en el precio, se observan las siguientes correlaciones:

Número de Dormitorios: Existe una correlación negativa entre el número de dormitorios y el precio de la casa. Esto significa que, a medida que aumenta el número de dormitorios, el precio de la vivienda tiende a disminuir.

Como la población está bien distribuida las conclusiones representan la realidad de las bases de datos.

Número de Baños: Se encuentra una correlación positiva entre el número de baños y el precio de la casa. Es decir, a medida que aumenta el número de baños, el precio de la vivienda también incrementa.

Podría ser que el precio se dispare a partir de 5 porque como las muestras son escasas pueden no representar a todas las casas con 5 baños.

Superficie Habitable (*sqft_living*): Hay una correlación positiva entre la superficie habitable y el precio de la vivienda. Así, a medida que aumenta la superficie habitable, el precio de la casa también aumenta.

Quizás los precios tan altos a partir de 6000 se deben a que la mayoría de estas casas sean mansiones, si es cierto que el precio igual no sería tan alto pues al haber pocas casas el precio no es significativo.

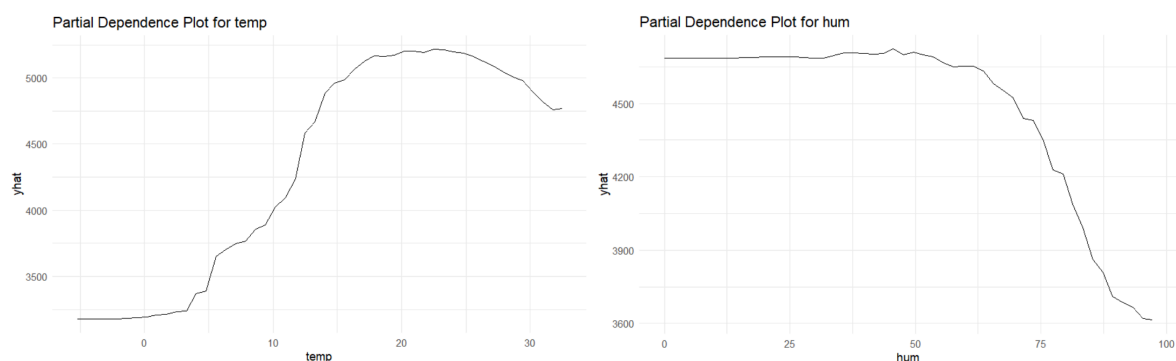
Número de Pisos: El análisis muestra una correlación positiva entre el número de pisos y el precio de la vivienda. Esto implica que, a medida que aumenta el número de pisos, el precio de la casa también incrementa.

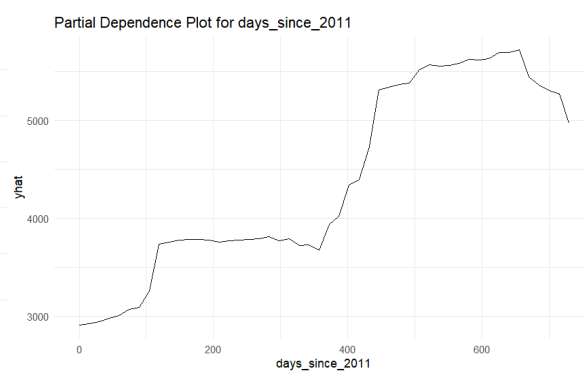
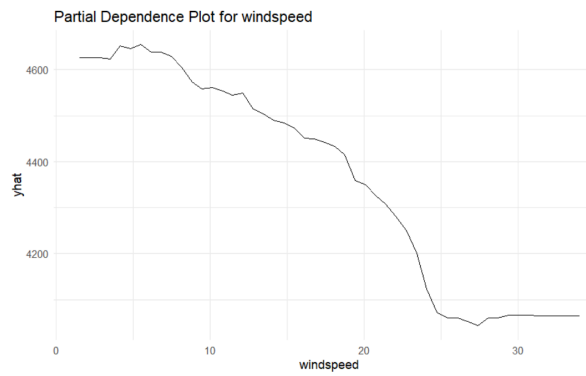
Al estar bien distribuida la población representa la naturaleza de la base de datos.

Estas correlaciones proporcionan una visión clara de cómo diferentes características estructurales de una casa influyen en su precio, ofreciendo información valiosa para la toma de decisiones en el mercado inmobiliario.

4. Anexo

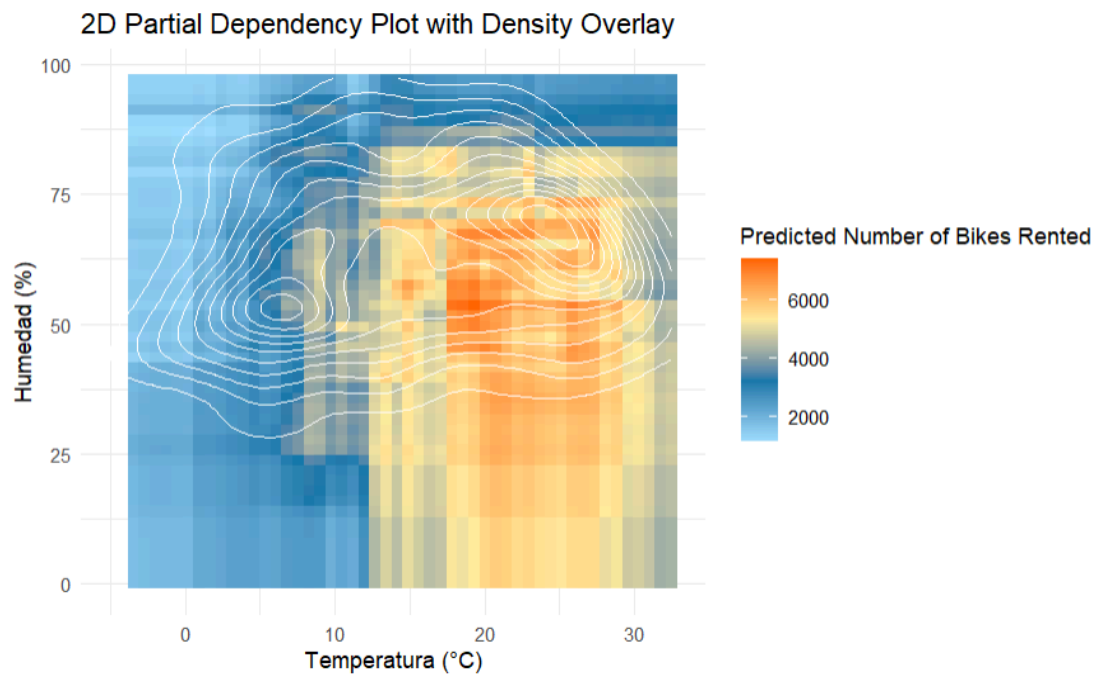
4.1 Gráfico dependencia parcial (bicicletas)





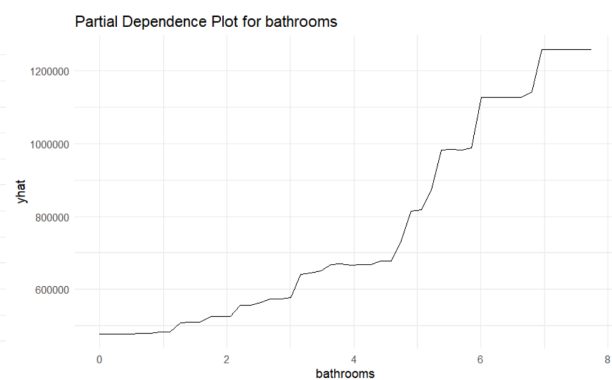
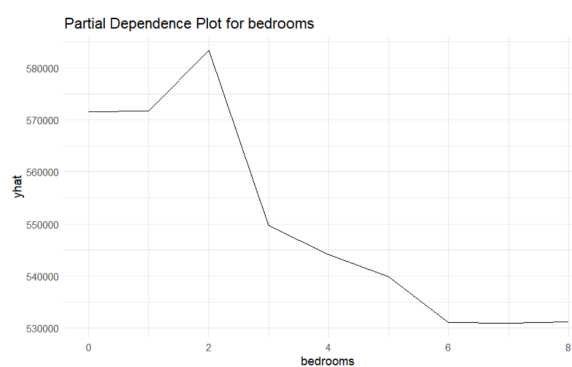
[Click para volver al informe](#)

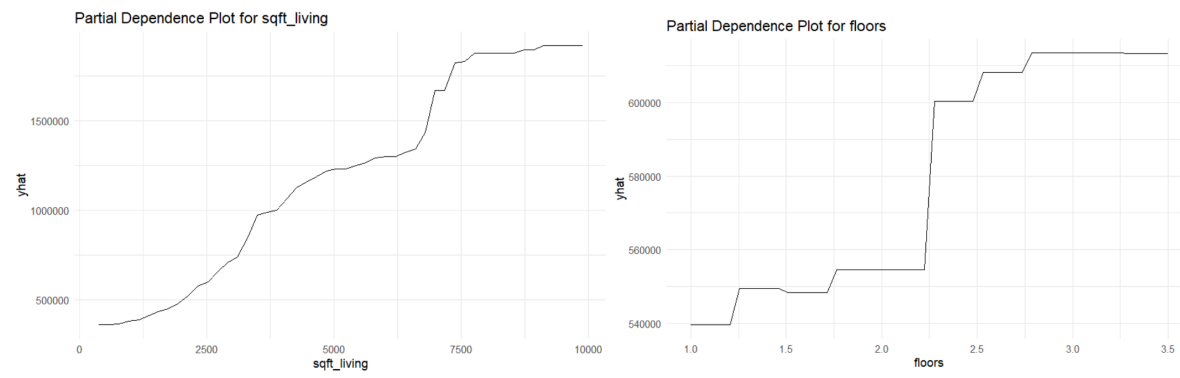
4.2 Gráfico dependencia parcial bidimensional



[Click para volver al informe](#)

4.3 Gráfico dependencia parcial (precio vivienda)





[Click para volver al informe](#)