MODEL-AGNOSTIC – PDP

Elena Chirivella Berto

EDM 2025 – GRADO EN CIENCIA DE DATOS UPV

Introduction

This project aims to apply model-agnostic interpretability techniques, specifically Partial Dependence Plots (PDPs), to analyze the behavior of Random Forest-based regression models. Throughout the exercise, a practical and systematic approach was followed, covering all key stages of the predictive analysis and interpretability process:

First, real-world data related to bicycle rentals and housing prices were loaded, cleaned, and sampled. Second, Random Forest regression models were trained, fine-tuned to predict the variable of interest (cnt in the case of bicycles, and price in the case of housing). In addition, one- and two-dimensional PDPs were generated, showing how different explanatory variables influence the model's prediction. Finally, a detailed analysis of the results was performed, drawing conclusions about the model's behaviour and the relationship between the variables and the target variable. In the first block of the project, a model is trained to estimate the number of daily bicycle rentals based on time, temperature, humidity, and wind, visualizing both the individual influence of each variable and the interaction between temperature and humidity. In the second block, a model is trained on a real estate dataset to predict home prices based on characteristics such as the number of bedrooms, bathrooms, square footage, and number of floors.

The entire process was developed in R, using specific libraries such as randomForest, pdp, ggplot2, and dplyr, and follows reproducible logic that allows black-box models to be interpreted from a comprehensible and visual perspective.

1. PDP unidimensional – Bike Rentals

In this first part of the project, the influence of different explanatory variables on the number of bicycles rented (cnt) is analyzed using one-dimensional partial dependence plots (PDPs). To do this, a Random Forest regression model was trained using the daily bicycle rental dataset.

The model was fitted with the following predictor variables:

• instant: number of days since January 1, 2011.

• temp: normalized temperature (values between 0 and 1).

• hum: humidity level (also normalized).

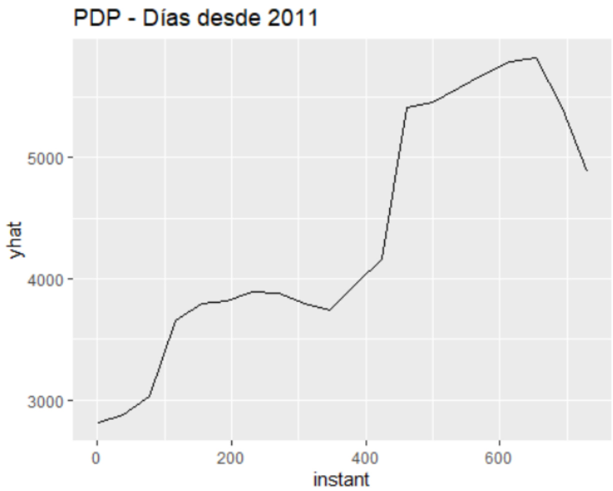
• windspeed: wind speed.

Once the model was trained, the PDP was calculated for each of these variables to visualize their marginal effect on the average prediction of the number of rentals, holding the other variables constant.

The analysis for each variable is presented below:

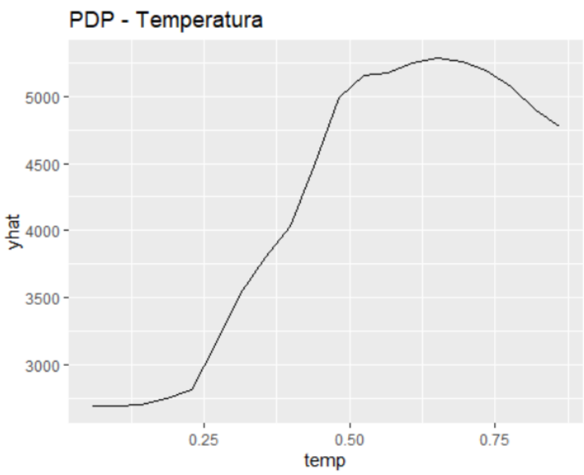
1.1. Instant (días desde 2011)

The graph shows a slight upward trend in the number of bicycles rented over the course of the day since 2011. This trend suggests a positive trend in the use of the rental service, possibly linked to greater implementation of the system, improvements in cycling infrastructure, or a change in urban mobility habits.



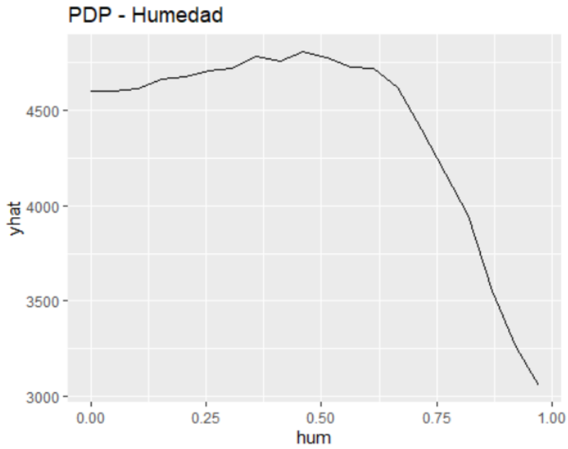
1.2. Temp (temperatura)

In this case, the graph shows a clearly positive relationship between temperature and the number of rentals. As temperature increases, so does the number of bicycles rented. The increase is not linear: the sharpest increase is observed with average temperature values. This result is consistent with the expected behavior, as good weather generally favors the use of bicycles as a means of transportation or leisure activity.



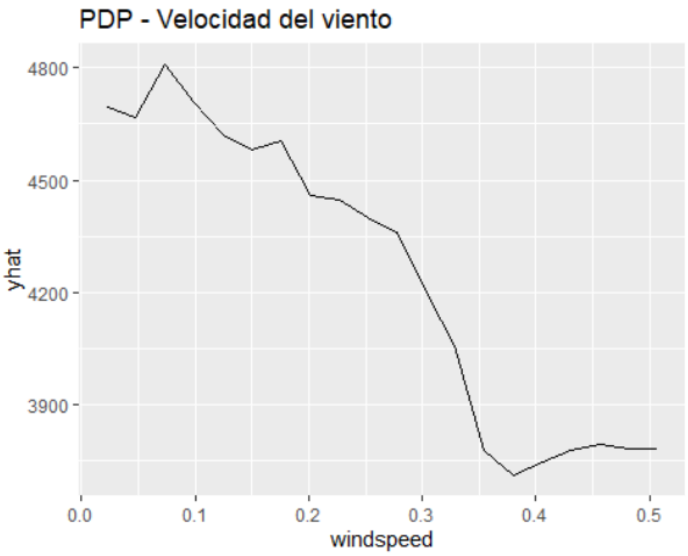
1.3. Hum (humedad)

The relationship between humidity and the number of rentals is inverse. As relative humidity increases, the model's prediction indicates a progressive decrease in the number of rented bicycles. This can be interpreted as a preference among users to avoid traveling in high humidity conditions, which are often associated with discomfort or bad weather.



1.4. Windspeed (velocidad del viento)

Finally, the graph for wind speed also shows a negative relationship. In high wind conditions, the number of rentals decreases. This pattern makes sense from a practical standpoint, as cycling in strong winds is more difficult and less comfortable, which can discourage users.



Taken together, these graphs clearly identify the variables that most influence the model's prediction. According to the importance calculated during Random Forest training, the variable instant is the most influential, followed by temp, hum, and finally windspeed. However, from an interpretive perspective, temp stands out as one of the variables most clearly associated with an increase in rents, while hum and windspeed are related to reductions.

2. PDP bidimensional – Bike Rentals

In this second section, the joint interaction of two explanatory variables on the prediction of the number of bicycles rented (cnt) is explored using a two-dimensional partial dependence plot (2D PDP).

Specifically, a PDP plot was generated using the variables temp (temperature) and hum (humidity), which showed significant effects individually in the previous section. The objective is to analyze how the combination of both variables affects the output of the Random Forest model, identifying joint patterns that are not apparent in one-dimensional PDPs.

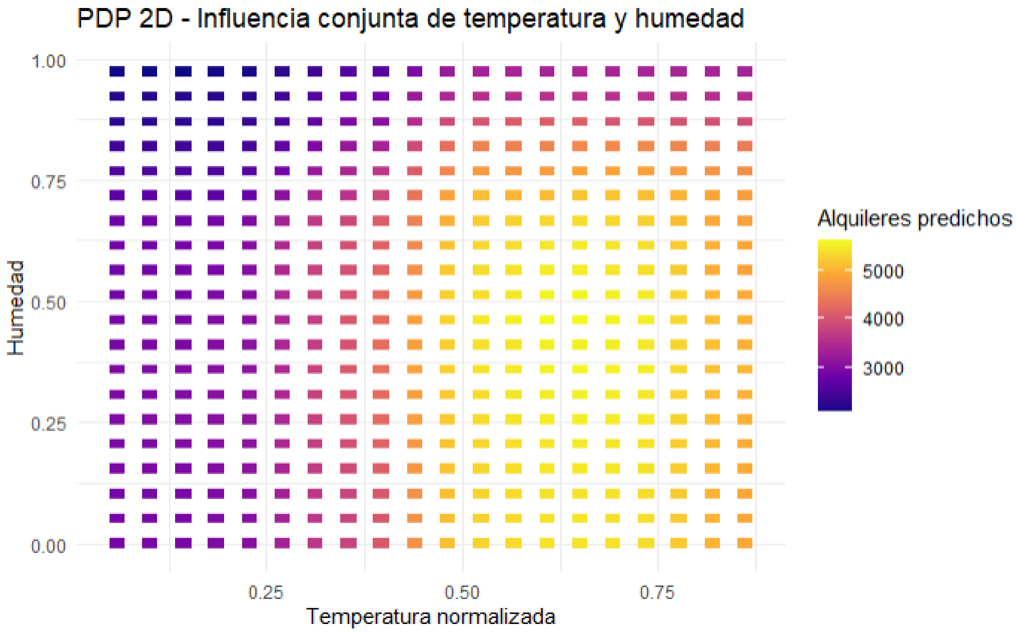
Given the size of the original dataset, random sampling was applied to reduce the computational burden without compromising the quality of the analysis. The two-dimensional PDP was then generated using the partial() function of the pdp package and visually represented as a heat map with geom_tile(), adjusting the width and height of each cell to avoid gaps.

The resulting heat map represents the normalized temperature values (temp) on the X axis and the humidity values (hum) on the Y axis. The color indicates the number of bicycles rented predicted by the model (yhat), where more intense tones correspond to a higher number of rentals.

The results show that optimal conditions for a high number of rentals occur when the temperature is high (above average) and humidity is moderate or low. However, in scenarios where the temperature is low or humidity is very high, the rental prediction decreases significantly.

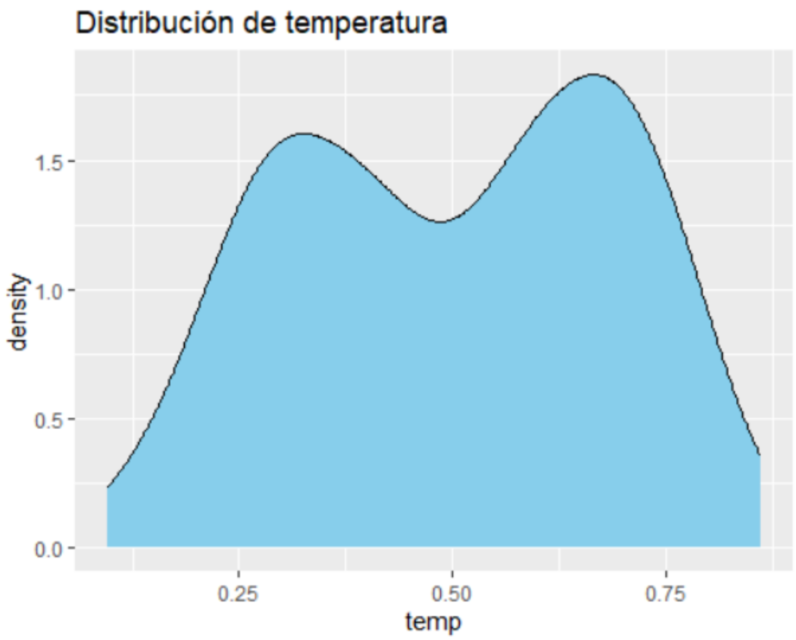
This behaviour is consistent with common sense: pleasant temperatures and a dry environment encourage bicycle use, while cold temperatures or high humidity levels can act as deterrents.

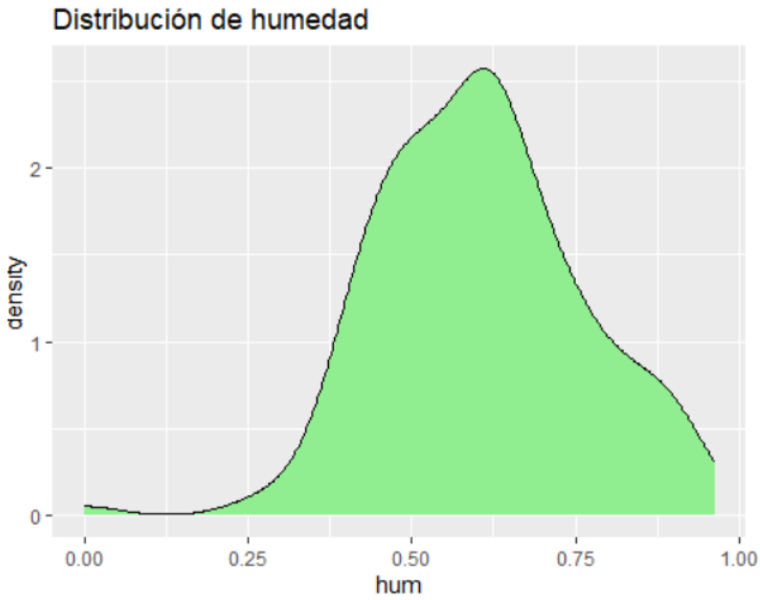
Furthermore, the analysis confirms that temperature has a more dominant effect than humidity. Although both factors play a role, differences in the number of predicted rentals are due more to changes in temperature than to variations in humidity.



To complement the interpretation, density distributions for temp and humidity were generated on a random sample of the dataset. These graphs allow us to visualize the frequency with which certain temperature and humidity values appear in the data, which helps contextualize the information shown in the heat map.

The density of temperature shows a greater concentration in average values, while humidity has a more dispersed distribution, with a slight bias toward high values. This implies that although ideal rental conditions occur with low humidity, they are not always the most frequent in the data.





3. PDP unidimensional – Precio de vivienda

In this section, the one-dimensional partial dependence plot (PDP) technique is applied to the problem of house price prediction using a Random Forest regression model. The objective is to interpret the individual influence of different structural characteristics on the value estimated by the model.

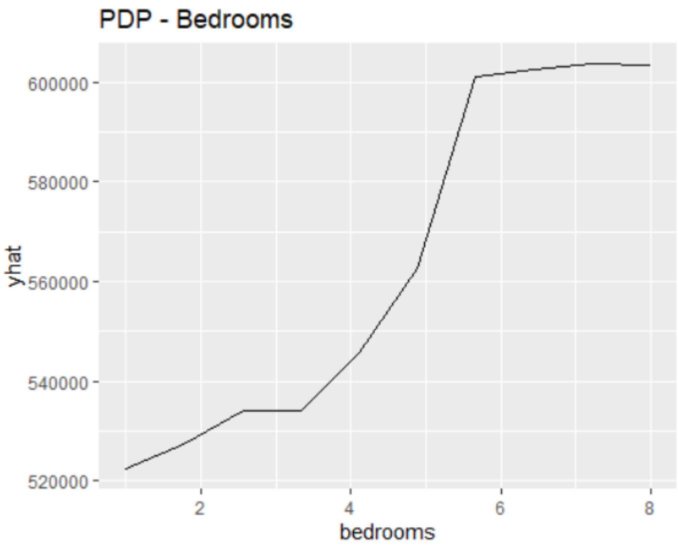
The model was trained on a random sample of the kc_house_data.csv dataset, with the following explanatory variables:

- bedrooms: number of bedrooms.
- bathrooms: number of bathrooms.
- sqft_living: usable square footage of the house.
- sqft_lot: total land area.
- floors: number of floors of the house.
- yr_built: year of construction.

For this section of the analysis, one-dimensional PDPs were generated for the four most relevant variables to understand how each impacts the price predicted by the model.

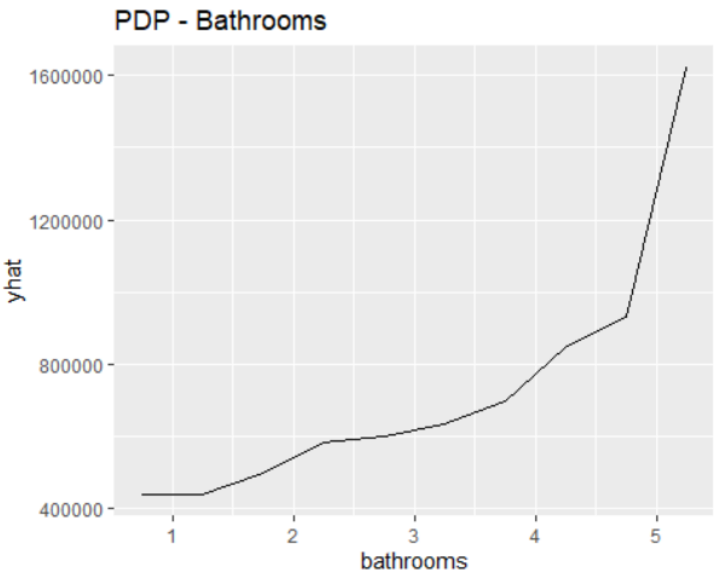
3.1. Bedrooms (número de habitaciones)

The graph shows that the effect of this variable on the predicted price is nonlinear. Up to approximately 4 or 5 bedrooms, the price increases slightly. However, beyond that point, the relationship stabilizes and even decreases for some values. This indicates that a greater number of bedrooms does not necessarily imply a higher price, possibly because many homes with many bedrooms correspond to less desirable market segments or have limitations in other aspects (location, quality, etc.).



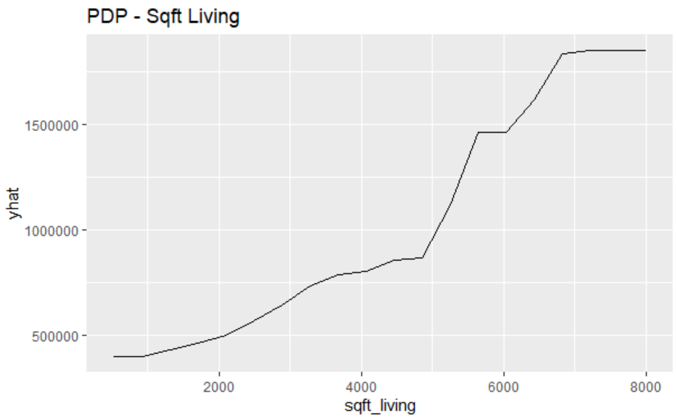
3.2. Bathrooms (número de baños)

The number of bathrooms is positively related to price. As this variable increases, the model predicts higher prices. However, the effect also levels off above a certain threshold (around 3 or 3.5 bathrooms). This result is reasonable, as having more than one bathroom is valued positively by buyers, although beyond a certain number it does not provide significant additional value.



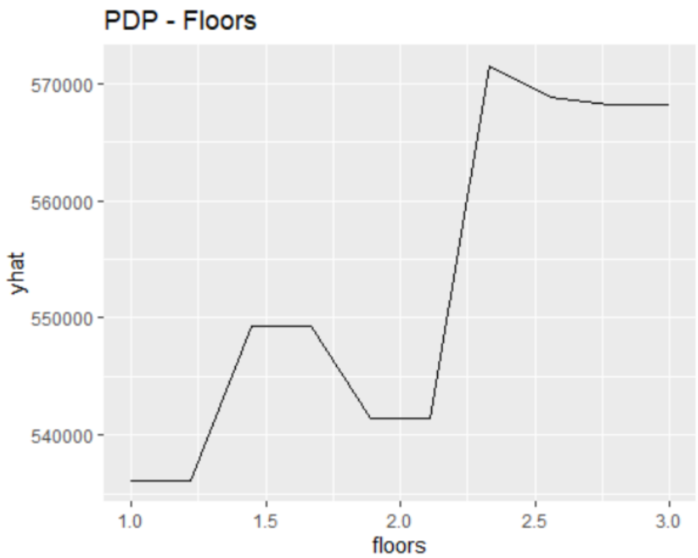
3.3. Sqft_living (superficie útil)

This is, without a doubt, one of the variables with the greatest impact on the predicted price. The graph shows a strongly increasing relationship: the larger the living area, the higher the price estimated by the model. The increase is almost linear, especially in the mid-range, reflecting that available interior space is a fundamental determinant of a home's value.



3.4. Floors (número de plantas)

In this case, the graph reveals that two-story homes tend to have higher values than single-story homes. However, the effect of the number of floors is not as pronounced as in the case of floor area. It is also observed that having more than two floors does not necessarily translate into higher prices, suggesting that this factor contributes to some degree of value.



Taken together, these results allow us to identify which characteristics have the greatest impact on predicting a home's price. Of all the variables analyzed, sqft_living stands out as the most decisive, followed by bathrooms. The number of bedrooms and floors also have an influence, but with more limited or nonlinear effects.

This type of analysis provides useful information both for understanding the model's behavior and for supporting decisions in real estate contexts, helping to estimate a property's value based on its key characteristics.

Conclusion

Throughout this project, the usefulness of partial dependence plots (PDPs) has been demonstrated as a fundamental tool for the interpretability of complex predictive models, in this case Random Forest regression models.

By applying them to two real-life cases—predicting the number of rented bicycles and estimating home prices—it has been possible to clearly and accessibly visualize the relationship between the input variables and the predictions generated by the model, both individually and in the interaction between two variables.

One-dimensional PDPs have allowed us to identify key patterns, such as the positive effect of temperature on rents or the increasing influence of usable area on home prices. Meanwhile, the two-dimensional PDP has provided a richer perspective by showing how the combination of temperature and humidity jointly affects bicycle service use.

Beyond the specific results, this work highlights the role of model-agnostic techniques in opening the black box of machine learning algorithms, facilitating transparent, visual, and understandable interpretation. In a context where trust, traceability, and explainability are increasingly necessary, PDPs are consolidating as an indispensable resource for translating complex predictions into useful conclusions for decision-making.