XAI 3 Report

MODEL-AGNOSTIC:
PARTIAL DEPENDENCY PLOT

Marián García Mas
Violeta García Rodríguez

INDEX

1. Introduction

Understanding the **influence** of individual factors on predictive outcomes is a critical aspect of data analysis and modeling. Model-agnostic methods, such as Partial Dependence Plots (PDPs), offer powerful tools to visualize and interpret the relationships learned by complex models. By isolating the effect of each feature, these graphs help in comprehending how changes in specific inputs impact the predicted results, making these methods extremely useful for transparent and interpretable machine learning.

This report explores the application of PDPs in two distinct but equally significant contexts: **bike rentals** and **housing prices**. Bike-sharing systems generate extensive data that can reveal insights into urban mobility patterns. Analyzing this data with PDPs allows us to understand the factors influencing bike rentals, enhancing our knowledge of these systems.

Similarly, in the real estate market, predicting housing prices accurately is extremely important for buyers and sellers. By applying PDPs to housing price predictions, we can dissect the influence of various features on the model's outcomes, clarifying what affects property values.

2. One dimensional PDP

Firstly, we carry out a comprehensive analysis of urban bike rental patterns using PDPs with four key variables—temperature, days since 2011, humidity, and wind speed— which, according to previous analysis, have been the most influential features. These plots illustrate their individual impacts on predicted bike rentals and also include "rug" marks, which indicate the density of data points to show the distribution of actual values that influence the model's predictions.

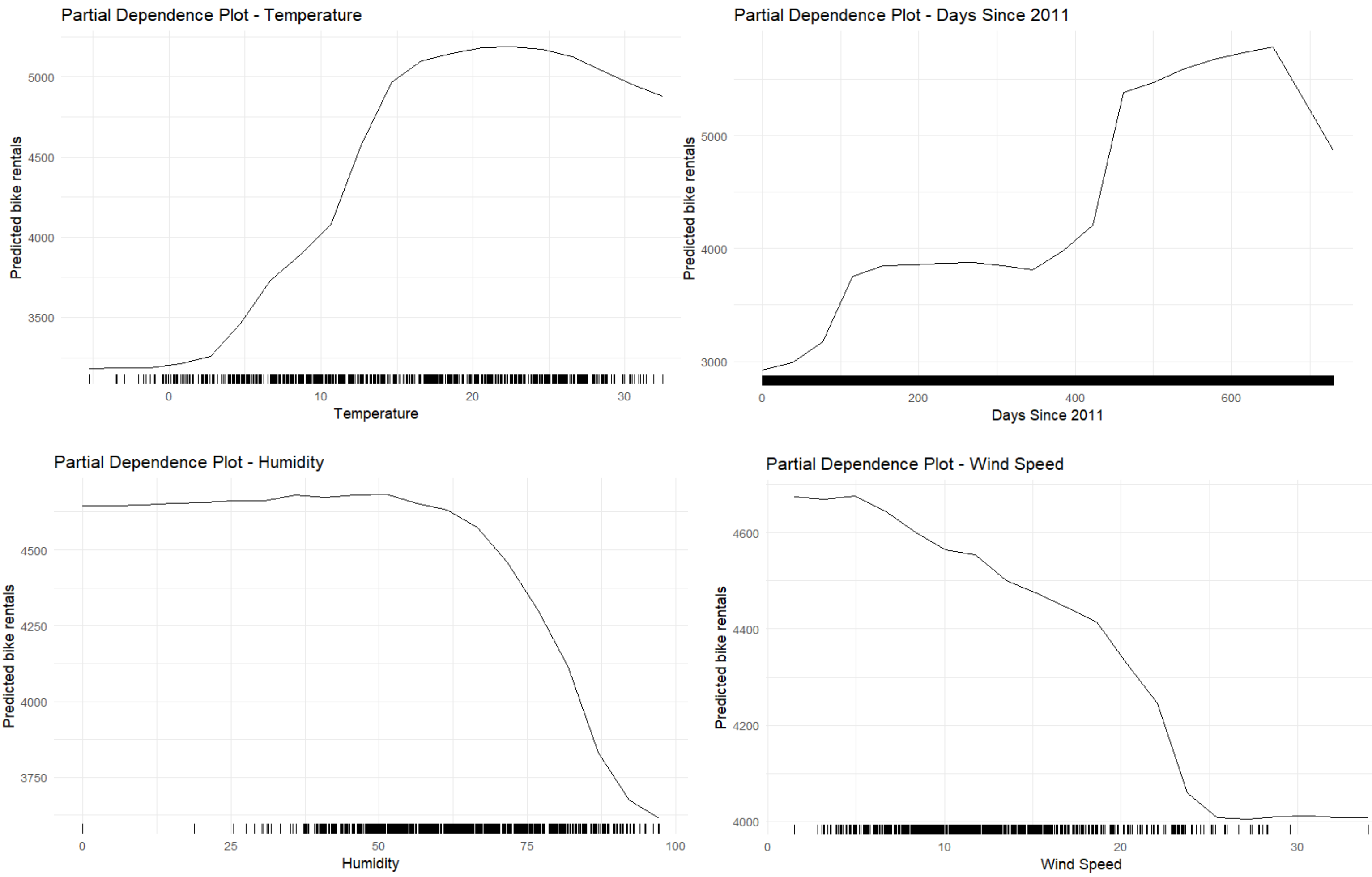

Figure 2.1: PDP Bike rentals vs. Wind Speed / Temperature / Humidity / Days_since_2011

The relationship between **temperature** and predicted bike rentals is characterized by a distinct bell curve, with a bit of negative skew, indicating an optimal temperature range for bike rentals. The plot shows that, as temperatures rise from sub-zero levels up to around 20ºC, the predicted bike rentals increase sharply, suggesting higher demand in better weather conditions. The peak in rentals occurs at about 30ºC, after which there's a slight decrease, likely due to discomfort from excessive heat. There is clearly a high density of observations between 10 and 30ºC, which ensures reliable predictions for the most typical temperature conditions. This pattern underscores the preference for cycling in comfortable, not overly hot weather.

The plot for **days since 2011** reveals two steep increases in bike rentals. The first one is from around 0 to 100 days after the bike-sharing program is introduced, suggesting a rapid and good reception of the bike-sharing system early on. Over time, the growth in rentals stabilizes and even shows minor declines around 400 days since 2011, which starts to be worrying. However, the second steep increase in bike rentals comes shortly after this low point - which could be due to the company introducing new incentives or an improved system. This increase in bike rentals continues on for about 200 days to a high point of around 6500 rentals, and then once again begins to fall. Overall, this gradual increase over the years may reflect growing societal trends towards healthier lifestyles or environmentally friendly transport options.
Moreover, the uniform distribution and the **fullness** of density bars indicates the plot's completeness, with daily recordings up to the present day and minimal exceptions, which underscores the remarkable credibility of the trends discovered.

The PDP for **humidity** shows a clear negative correlation with bike rentals. Starting from low humidity levels, there is a significant and consistent drop in the number of predicted bike rentals as humidity increases. The plot starts with higher predicted rentals at lower humidity, gradually decreasing as humidity rises, which highlights the adverse effects of moist conditions on cycling comfort. The density is higher at lower humidity levels, indicating a larger amount of data and thus more reliable predictions under these conditions.

Analyzing the **wind speed** plot, a clear negative relationship is evident between wind speed and bike rentals. The graph begins at around 4700 predicted rentals with nearly no wind, which shows the strong preference of people for cycling on days without wind versus days with wind. As wind speed increases, there's a gradual decrease in rentals, becoming sharply pronounced beyond 10 km/h, and even more so around 20, where the rentals plummet to approximately 4100. The sparse density at higher wind speeds suggests the lack of information on these types of days, due to very windy days being much less frequent , which makes potential predictions less certain in these ranges.

3. Bidimensional PDP

2D Partial Dependency Plots help in visualizing the **marginal effect** of a subset of features on the predicted outcome of a machine learning model. By focusing on temperature and humidity, we can interpret how these two environmental factors influence bike rental patterns, assuming all other variables remain constant.

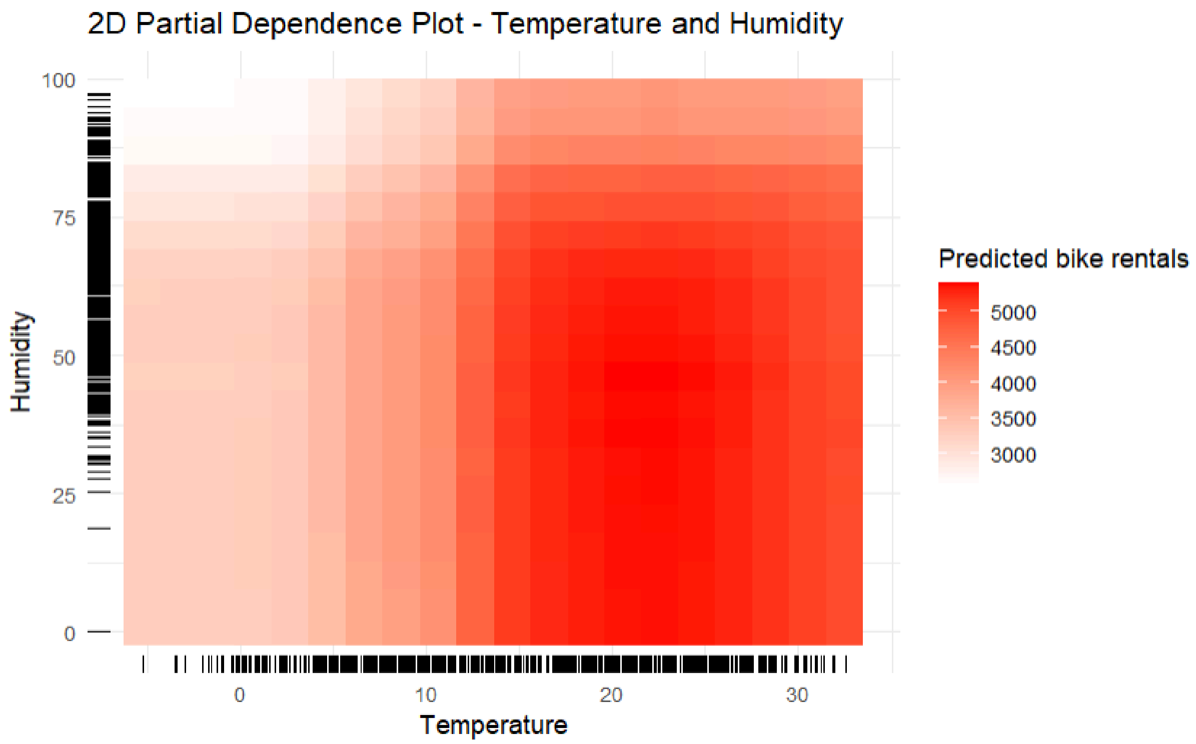

Figure 3.1: 2D PDP Temp vs. Hum

The pattern that appears in the plot is clearly visible; the intensity of bike rentals increases substantially as temperature **rises**, reaching its peak at the highest temperature range displayed. This suggests that higher temperatures are strongly associated with increased bike rentals, up to 30 degrees Celsius. This reflects a preference for biking in warmer weather, as seen in previous analysis, due to more enjoyable riding conditions or longer daylight hours that come with seasons with warmer temperatures like spring or summer.

Interestingly, the impact of humidity appears less significant in influencing rental numbers compared to temperature. Throughout the range of temperatures, the variations in humidity show only slight differences in rental predictions, with a gentle **decrease** in rentals as humidity increases, particularly noticeable at the highest temperature values. This could indicate that while riders prefer warmer days, they tend to want to use a bike less on extremely humid days, possibly because of the discomfort caused by the sticky weather conditions.

The density indicators along the axes of the plot, which are the frequency of temperature and humidity observations, are extremely important to consider when drawing conclusions from PDPs. A **denser** collection of data points at certain temperatures or humidity levels means that conclusions drawn from these areas are based on more substantial evidence and are more reliable. On the other hand, sparse data points in other areas suggest that conclusions should be interpreted with caution as they may be less certain.

For example, the information derived from the effect of temperature is fairly **reliable** as there is a good number of observations across most temperature ranges. There are fewer observations above 30°C and below 0°C, which is expected due to the rarity of these extreme temperatures. However, the conclusions drawn previously are still reliable because the observed pattern is primarily seen within the temperature range with the most amount of observations.

In contrast, the distribution of humidity observations is less evenly spread. Most observations are concentrated between 25% and 100% humidity, so making conclusions within this range, like we did before, is reasonably **safe**. There are very few observations of days with lower humidity levels, so we cannot reliably assume the bike rental behavior on these days.

4. PDP to explain the price of a house

The Partial Dependence Plots for the **housing market** provide a clear visualization of how certain features affect the predicted price of houses. These plots are obtained from a Random Forest model trained on a sample of housing data, focusing on four specific features.

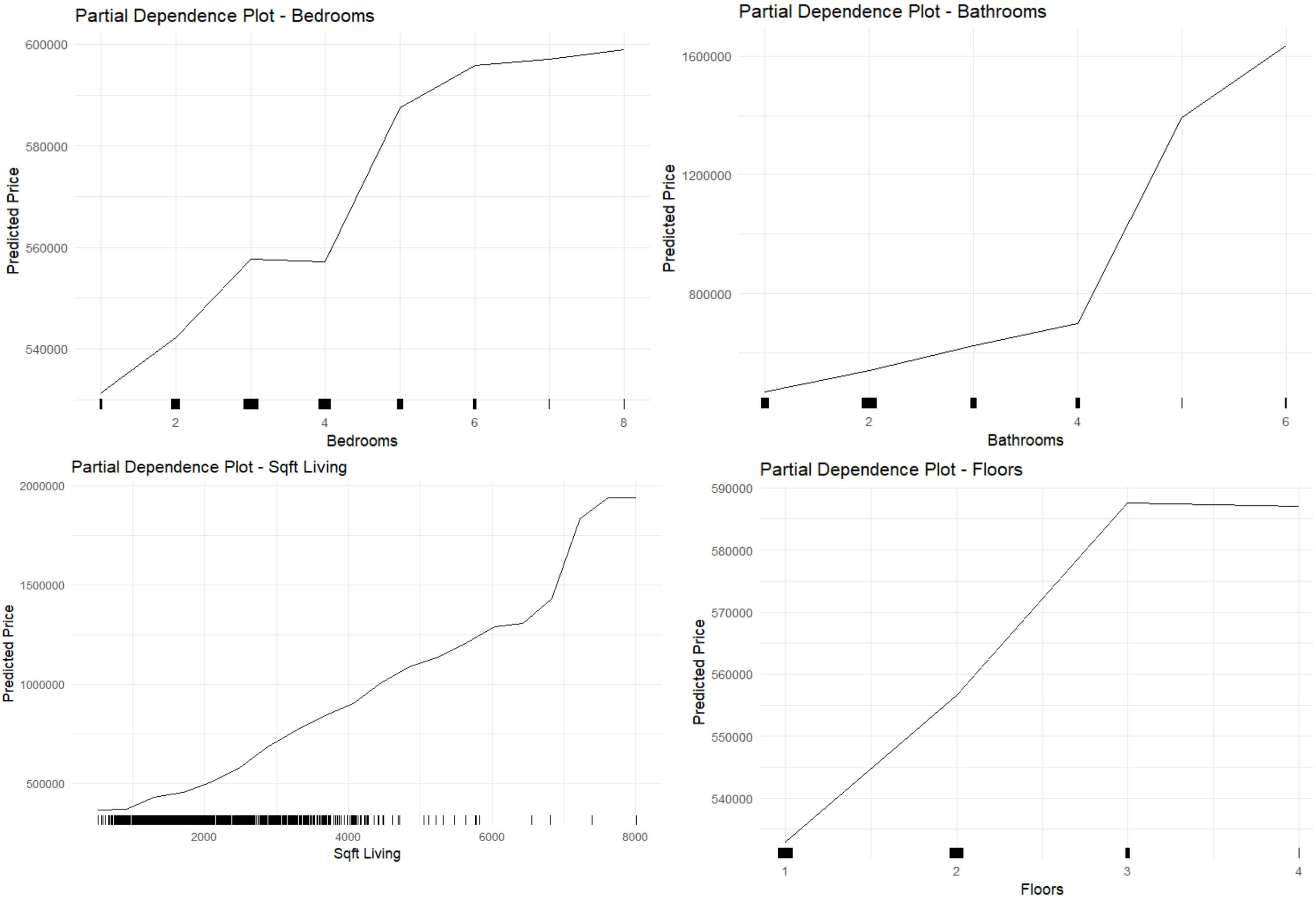

Figure 2.1: PDP Predicted price vs. Bedrooms/Floors/Sq Ft Living/Floors

The PDP for **bedrooms** shows a generally increasing trend in predicted prices as the number of bedrooms rises from one to four. However, after four bedrooms, the increase in price stabilizes and even slightly declines at eight bedrooms. This may be because of a practical limit to the value added by additional bedrooms in a typical family home, as these houses may be less appealing to the average home buyer. The density marks, which become thicker the higher the amount of observations there are for each category, show a high concentration of data points for homes with fewer bedrooms (up to four), making the predictions more reliable in this range. The thinner line of marks at higher bedroom counts indicates fewer observations, indicating less reliability in these predictions.

The plot for **bathrooms** displays a gradual increase in predicted prices as the number of bathrooms increases from one to four. Beyond four bathrooms, there is a steep price increase up to six bathrooms, nearly doubling from around $800,000 to $1,600,000. However, the very thin lines of density marks at six and seven bathrooms indicate minimal data points, which shows us that conclusions about the extremely high prices associated with more bathrooms are not reliable at all, and so we can not assume anything about houses with this configuration.

The plot for **square footage of living space** illustrates a strong positive correlation with predicted house prices. As the living area increases, so does the price at a fairly constant rate. The data density is highest in the 1,000 to around 4,000 square foot range, meaning these are common house sizes, and so we have more information about them. Although it is common sense that the bigger houses + expensive… we cannot assume anything bc of low amount of data….

The PDP for **square footage of living space** indicates a strong positive correlation with house prices. As the living area increases, so does the predicted price. The density of data is highest between 1,000 and 4,000 square feet, confirming that these are typical house sizes with ample data, and so we have more information about them. Although it is common sense that the bigger houses, the more expensive it will be, we can not assume anything once again because of the low amount of data of houses with areas between 6000 and 8000 square feet.

The plot related to the **number of floors** shows a clear increasing trend in predicted prices from one to three floors, suggesting a preference for multi-story homes, potentially due to the perception of more space or separation between living areas. The price stabilizes from three to four floors. The density marks are thicker for homes with one to three floors, indicating a good amount of data and reliable predictions for these categories, but, like before, fewer observations at higher number of floors (four) make the price prediction at this level less certain.

5. Conclusion

We can finally conclude that the application of Partial Dependence Plots in the analysis of both datasets has provided valuable insights into how changes in specific variables **impact** the predictions of complex machine learning models.

In the analysis of bike rental patterns, we observed notable effects of temperature, days since the start of the bike-sharing program, wind speed, and humidity on predicted rentals. The duration of the bike-sharing program showed periods of rapid growth and loss of interest in the program, but increased overall, reflecting the increasing popularity of biking over time. Temperature was also positively correlated with rentals, indicating a preference for biking in comfortable weather. In contrast, wind speed and humidity showed negative correlations, showing disinterest in biking in uncomfortable conditions. These relationships are relatively **reliable** due to the ample amount of data available in nearly all ranges of the variables.

In the housing market analysis, we found that most features **positively** influenced predicted house prices. The number of bedrooms, bathrooms, and square footage of living space all showed positive correlations with prices, aligning with the common understanding that larger and more feature-rich homes tend to be more valuable. However, reliable conclusions could not be made for these so-called “feature-rich homes” because of the lack in number of these type of houses, and so lack of observations.

Additionally, the analysis highlighted the significance of considering the density of observations when interpreting PDPs. In regions with **limited data**, predictions become less certain, emphasizing the need for caution and careful interpretation.