XAI 3 Model-agnostic methods
Partial Dependency Plot

Authors:
Emilie Sivertsen and Nathalia Thu

Date: May 15, 2025

Contents

Introduction

In this report we will interpret the results from plotting a Partial dependency plot (PDP) on the bike rental dataset. We will look at both one dimensional and bidimensional.

We will also plot some PDP on a house pricing dataset, to be able to see how the features influences the pricing.

A PDP plot is used on a trained model where we change the values of one feature to see how it affects the model. This method assumes that the features are independent.

The code is located in github, ensuring version-control with git and backup support in github.

1 One Dimensional Partial Dependence Plot

To interpret some PDP plots we first have created a Random Forest Regressor on the bike rental dataset. A PDP plot for four different features in the bike rental model is shown in figure 1.

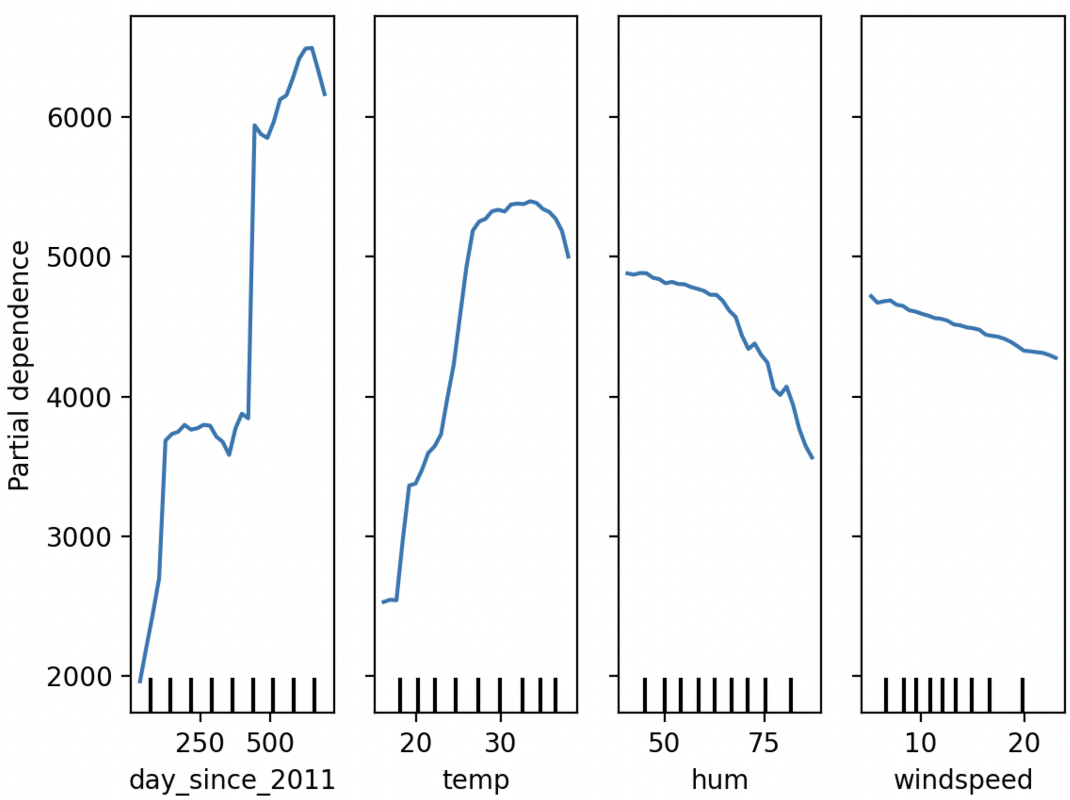

Figure 1: Partial Dependence Plot of days_since_2011, temperature, humidity and windspeed in the bike rental model.

Here we can see that both days_since_2011 and temperature in general increases with value, giving a higher rental bike rental count. While humidity and windspeed decreases, meaning it lowers the bike rentals.

When looking at days_since_2011 we see that it increases with time, but has some stops along the way. This can be interpreted as the seasons, where some seasons give more bike rentals than others. Meaning it increases by time in general, but because of seasonal dependences it has some stops along the way. From the temperature PDP we also see it increasing with increasing temperature, until the temperature gets too high - giving lower rental count. At the graph maxima we can see that it keeps the same count at around 25-35 degrees, which seems to be when most people rent bikes.

The humidity graph has an accelerated decrease where with a slow slope in the start which drops faster and faster. This means that high humidity values have a more impact on the result than low ones. The windspeed feature has a linear decreasing with increasing value.

2 Bidimensional Partial Dependency Plot

We can see even more complex correlation by plotting a PDP plot of multiple features. A plot of the bidimensional PDP of temperature and humidity is shown in figure 2.

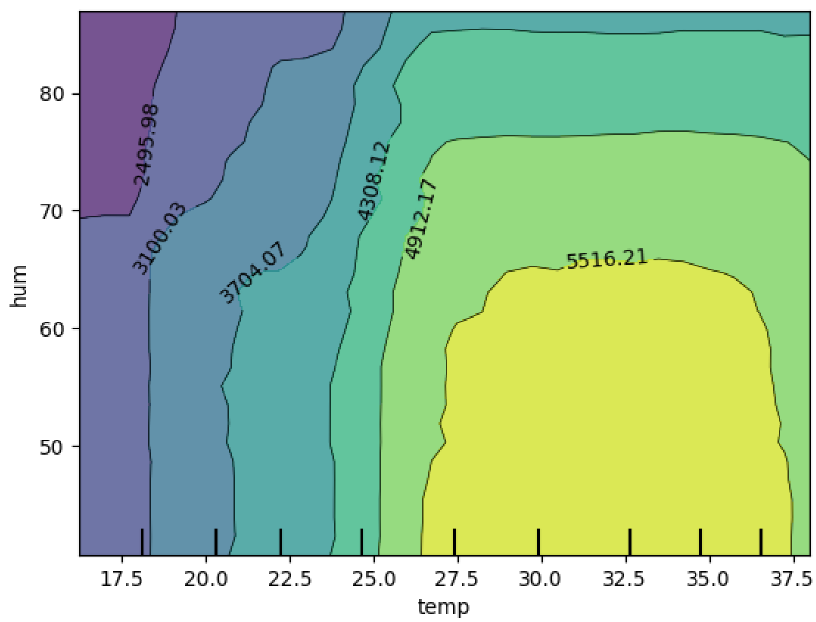

Figure 2: A 2D PDP for the temperature and humidity features in the bike rental model.

Here we can see that high temperatures and low humidity are the ideal weather for bike rentals (yellow area with highest bike rental values), while low temperatures and high humidity clearly have the lowest bike rentals(with darkest color).

We can also see that the lines are somewhat straight vertical when the humidity is low from 65 and down, meaning that here the number of bike rentals mainly depends on the temperature. The opposite can be said for temperatures between 26-37 where the lines are horisontal, meaning the humidity mainly determines the bike rental count.

3 PDP to explain the price of a house

Now we will use a new dataset, containing housing prices. We train a Random Forest Regressor on the features bedrooms, bathrooms, sqft_living, sqft_lot, floors, yr_built. Next, we want to know how the four features bedrooms, bathrooms, sqft_living and floors influence the target value price by interpreting PDP plots.

First we plotted a one dimensional PDP for each feature, this is shown in figure 3.

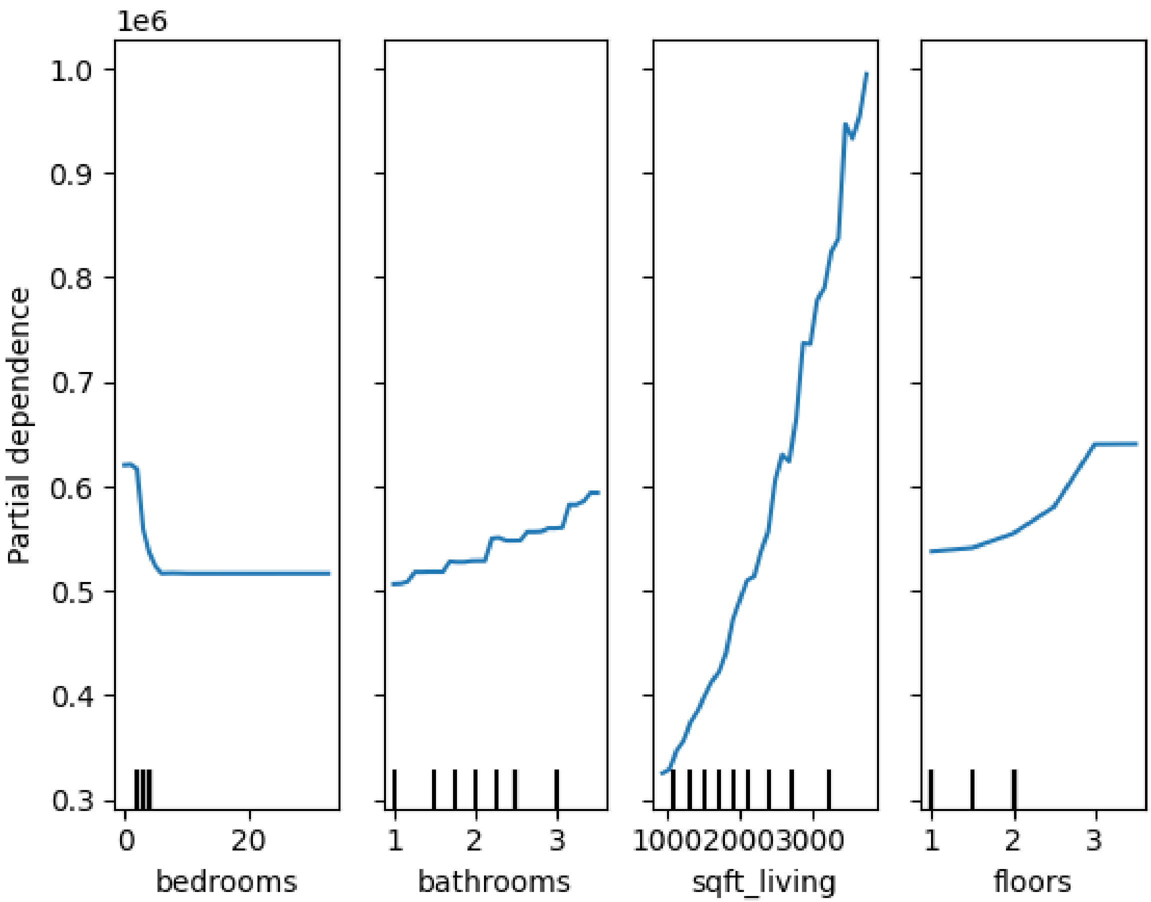

Figure 3: One dimensional PDP for the features bedrooms, bathrooms, sqft_living and floors in the house pricing model.

In the 1D PDP plot we see that it is a higher pricing for few bedrooms, while if the number of bedrooms go past 5 we have a somewhat constant lower pricing. This shows that

For bathrooms we can see that there is a positive relation, more bathrooms means higher price. This trend is relatively linear and smooth, indicating a consistent increase in predicted value with additional bathrooms.

The variable sqft_living has the strongest positive influence, where bigger living space means higher pricing. The curve is steep, giving a much better price for a slight change in space.

Also shows a positive correlation, but weaker than sqft_living or bathrooms. Most of the impact is from 1 to 2 floors, then it levels off or increases slightly.

Next, 2D PDP plots of the four features bedrooms, bathrooms, sqft_living and floors are shown in figure 4. The more colour (yellow) it is, the higher price.

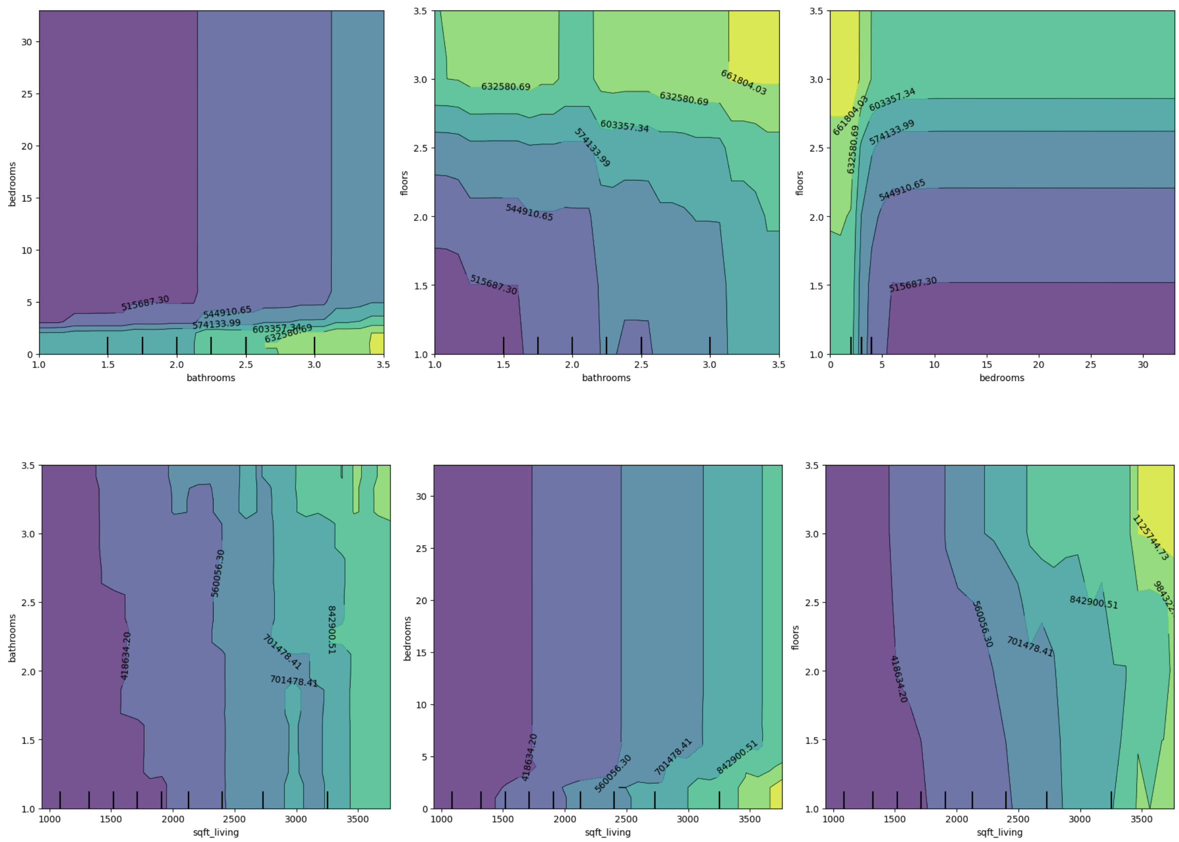

Figure 4: 2D PDP plots of theh features bedrooms, bathrooms, sqft_living and floors.

The 2D PDP plot for bathrooms and bedrooms shows that an increase in bathrooms leads to a higher predicted price, particularly up to about 3.5 bathrooms. However, the effect of bedrooms is minimal — increasing the number of bedrooms beyond 5 does not significantly change the predicted price. We can also see this in the plots with bedrooms/floors and bedroom/sqft_living, where low value for bedrooms gives the best pricing, and when there is many bedrooms the other value decides the pricing. This indicates that bathroom, sqft_living and floors are a stronger predictor of price than bedrooms.

In the bathrooms and floors plot, we observe a clear positive interaction. Houses with both more floors and more bathrooms tend to be valued higher by the model. For instance, prices rise notably when moving from one to two floors, especially when combined with 2.5 or more bathrooms. The combination of three floors and three bathrooms leads to some of the highest predicted prices. This suggests that bathrooms and floors are both valuable features when combined.

At the bottom row in the figure we have sqft_living on the x-axis. The contour lines are mostly vertical, indicating that living are is the dominant factor. However, both bathroom and floors increases the pricing value where higher pricing values are for more bathrooms and more floors. This indicates that bigger living space, more bathrooms and more floors increase housing value. But simply increasing the bedroom count does not lead to higher predicted values.

Conclusion

In this report, we applied Partial Dependence Plots (PDPs) to interpret how selected features influence predictions made by Random Forest models in two different datasets: bike rentals and housing prices.

For the bike rental dataset, the PDPs revealed that warmer temperatures and more recent dates (interpreted through days_since_2011) are positively correlated with higher bike rental counts. On the other hand, higher humidity and windspeed led to lower predicted rentals. The 2D PDP further illustrated how low humidity combined with high temperature provides the most favorable conditions for rentals.

From the partial dependence plots (PDP) applied to the housing price dataset, we conclude that the most influential features on predicted house prices are sqft_living, bathrooms, and floors. Among these, sqft_living showed the strongest positive relationship with price, confirming that larger living spaces significantly increase house value. Bathrooms also had a nearly linear increase in price with number, while floors had a moderate but still notable positive effect, especially when increasing from one to two floors.

In contrast, the number of bedrooms had little impact beyond five bedrooms. This is supported by both the 1D and 2D PDPs, where interactions between bedrooms and other features showed diminishing returns in price prediction. This suggests that increasing living space and improving home quality with more bathrooms and floors adds more value than simply increasing bedroom count.

Overall, PDPs served as a valuable model-agnostic tool to visualize and interpret how different input features affect prediction outcomes in both use cases.