XAI 3 – MODEL-AGNOSTIC METHODS

Carlos Soriano Adam, Pol Llobet Valls and Federico Romiti

UNIVERSITAT POLITECNICA DE VALENCIA

Content

Technical Report: Partial Dependence Plot for House Price Prediction

1. Objective

In this report, we have presented an application for **Partial Dependence Plots (PDPs).** This allows us to interpret the behavior of a Random Forest model in two different tasks. On one hand, estimating bike rental demand based on different features such as weather and time related. On the other hand, we have predicted house prices from property attributes. We are looking forward to exploring how individual features influence model predictions and evaluate whether those effects are aligned with domain expectations.

PDPs offer a practical approach that helps to interpret black-box models by quantifying how changes in features affect the predicted outcomes. This makes them valuable tools in explainable machine learning work.

Part I – Bike Rental Demand

2. Model Overview

During the first half of this work, we learned a Random Forest model to predict the number of bike rentals on a day from the day.csv data set. This data set has variables such as weather (temperature, humidity, windspeed), season, and date. After preprocessing and feature engineering (e.g., denormalization of weather variables, creation of categorical indicators), a model was trained on a subset of features selected.

3. Theoretical Context

Partial Dependence Plots (PDPs) are model-agnostic methods for visualizing the average marginal effect of a feature on the model's prediction. By keeping all other variables constant, PDPs consider the effect of a single input and plot the effect of its values on the predicted outcomes.

This is especially important when interpreting a Random Forest, which can consider non-linear interaction but can also obfuscate what drives predictions. In this case analysis, PDPs allow us to evaluate how environmental and temporal variables impact predicted rental counts.

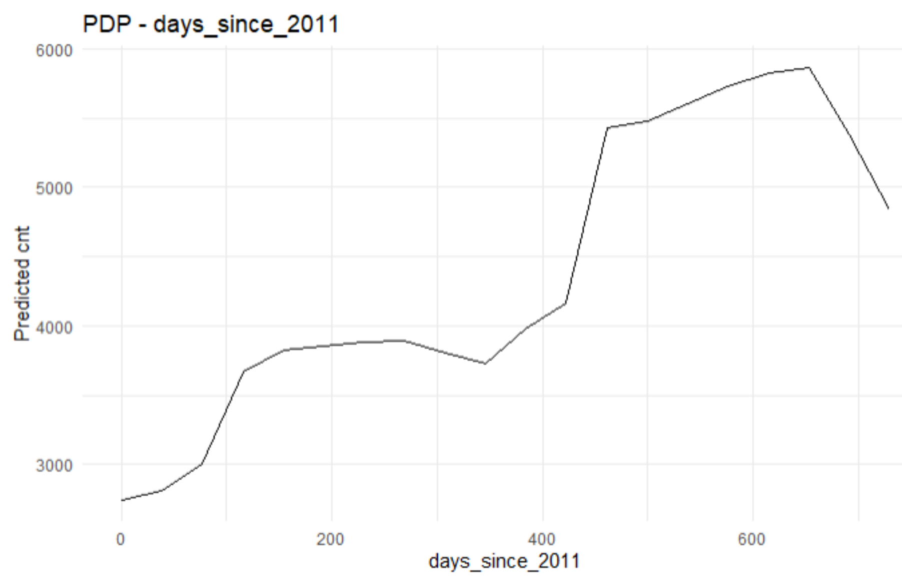
4. Feature Analysis through PDPs

The second segment includes a set of Partial Dependence Plots generated from the Random Forest model trained on the day.csv data. They come with an interpretation to see how marginal effects of key environmental and date variables—temperature, humidity, and date—on the predicted number of bike hires. These indicate if the model is learning relationships in line with actual human behavior and seasonality.

4.1. Temporal and Weather Effects

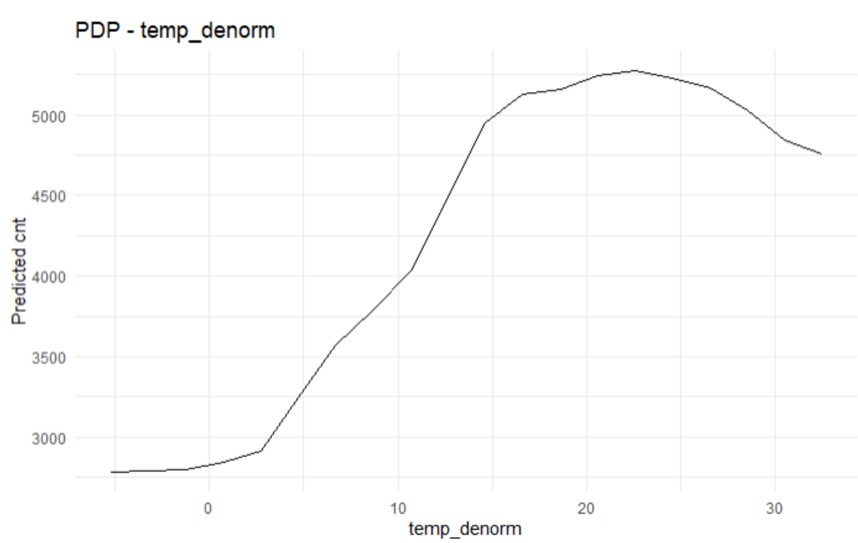
We begin by analyzing how time and weather-related features affect the model's predictions.

Days Since January 1, 2011



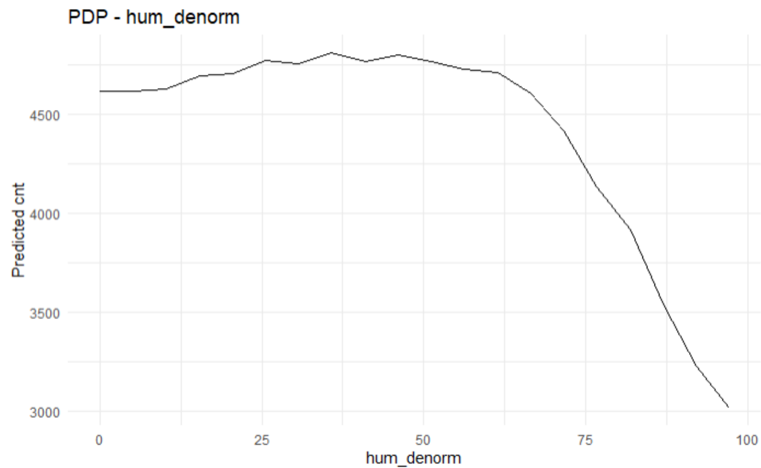
The number of days since the beginning of the dataset displays a general increase in predicted rentals over time. This likely indicates an increasing pattern in bike use over the period of interest, likely due to developing infrastructure, policy changes or simply a broader public adoption of bike-sharing systems.

Temperature (temp_denorm)



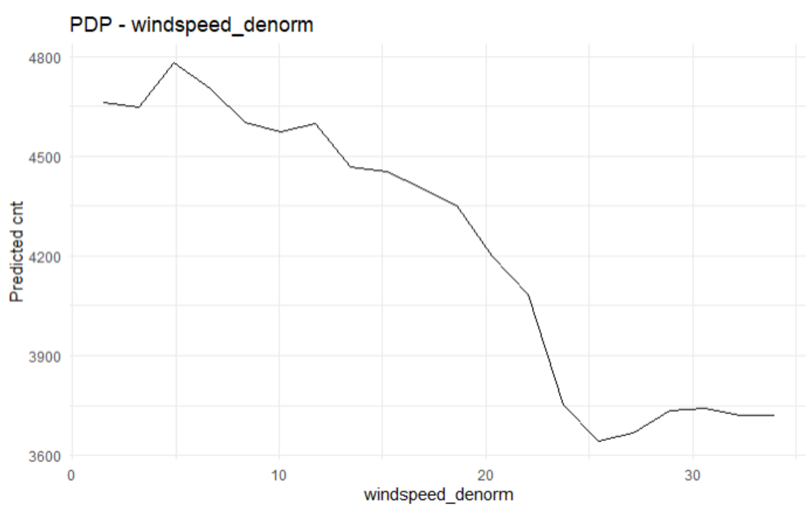
The PDP for temperature suggests a strong and positive impact on bike rentals, reaching a peak in effect size at around 30°C. This is in line with our understanding that warmer temperatures encourage outside activity to some degree. Beyond 30°C, the positive impact appears to plateau because warmer temperatures may cause discomfort.

Humidity (hum_denorm)



The PDP for humidity indicates a clear and negative relationship to the rental predictions. As humidity increases, the number of predicted rentals decreases, people are generally less likely to cycle under humid, sticky or perhaps mild rainy conditions.

Windspeed (windspeed_denorm)

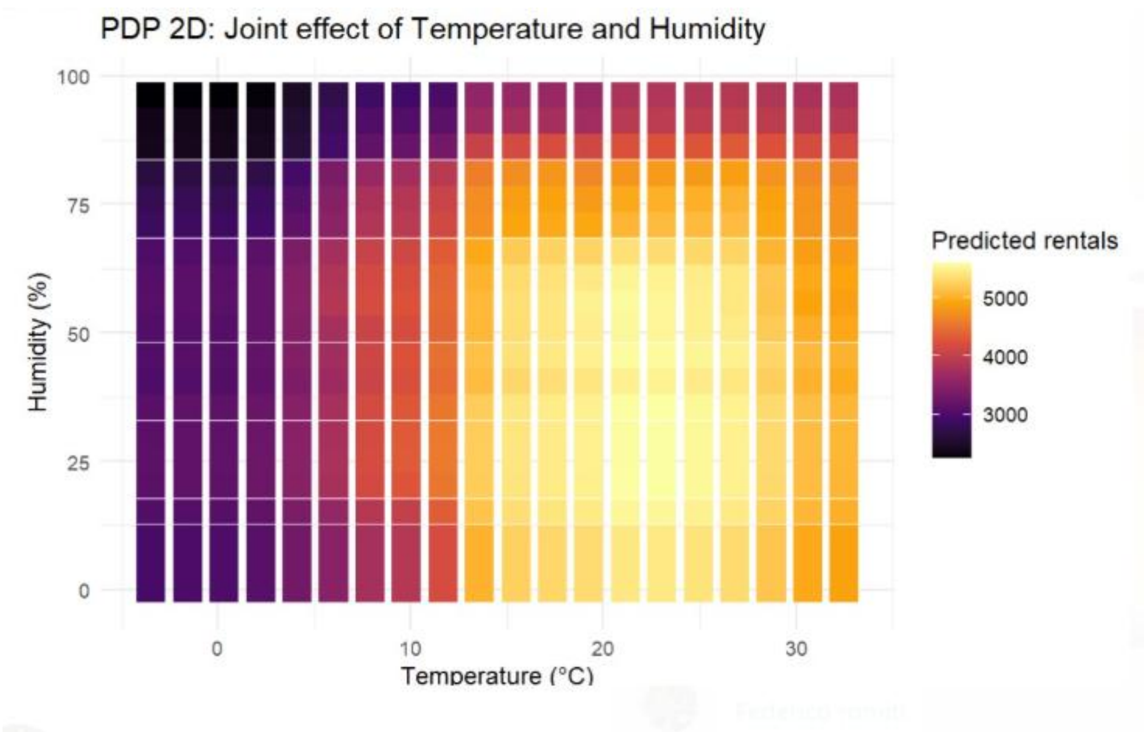


The PDP for windspeed exhibits a gently negative effect. As a non-motorized mode of transportation, cycling under moderate-to-high wind gusts appears less likely. Recent research has suggested that people are tolerant of light to moderate wind speeds when cycling: if windspeed increases, the probability of renting a bike decreases slightly, corresponding with the general trend of how people are known to behave.

4.2. Interaction Effects

In this section, we explore how temperature and humidity interact to influence predicted rental counts.

Temperature × Humidity (2D PDP)



The bi-directional PDP of temperature and humidity reveals their interaction effect. **Warm weather** with low humidity is linked to the highest predicted rental rates, and **high humidity** suppresses demand irrespective of temperature. This confirms the model's ability to capture non-linear interactions in weather-responsive behaviors.

5. Conclusion

The Partial Dependence Plots analysis for the bike rental dataset showed that the Random Forest model of the bike rental data accounted for the influence of environmental and temporal factors on user behavior. Like previous results with temperature, predicted rentals had a strong positive effect on temperature, and a negative effect for humidity and windspeed, consistent with human levels of comfort and a reasonable assumption of the real world. Furthermore, the overall trend towards an increase over time mirrors a larger behavioral trend towards riding as distinct from other modes of transportation.

Overall, the PDPs verified that the model learned realistic and interpretable patterns, including main effects, significant non-linearities, and interactions, confirming the utility of PDPs for evaluating whether black-box models are behaving according to domain logic when predicting behavioral data.

Part II – House Price Prediction

6. Model Overview

As noted in the second portion of this report, we trained a separate Random Forest model using a random sample of 1,000 properties from the kc_house_data.csv dataset. The model predicted house prices based on several significant features: the number of bedrooms, number of bathrooms, square footage, size of the lot, number of floors, and year of construction.

This dataset is an excellent opportunity to assess how structural and temporal property attributes influence price as learners in the model.

7. Theoretical Context

As in the previous section, Partial Dependence Plots were used to describe model behavior. Averaging out all other variables, PDPs show the isolated effect of individual variables on predicted house price. They show here if the model captures true market trends, e.g., size premiums, decreasing returns, or time effects.

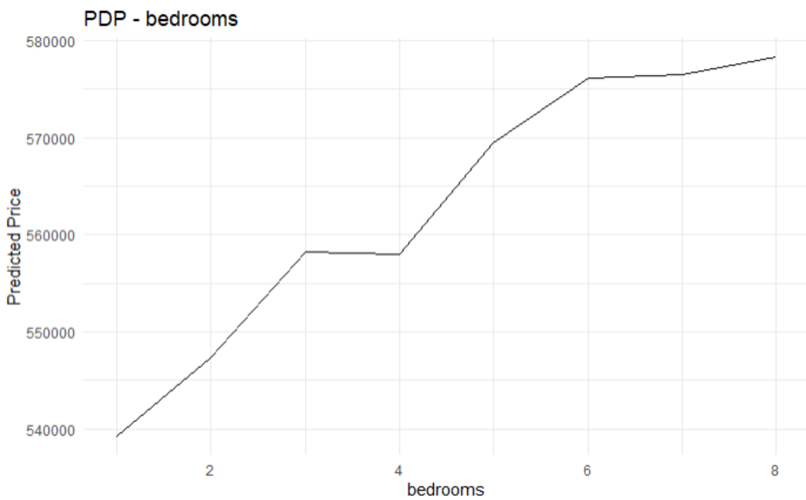
8. Feature Analysis through PDPs

This section demonstrates Partial Dependence Plots drawn from the Random Forest model that is trained to predict house prices. The plot displays marginal effect on a single input feature towards the predicted price and gives insights into how the inside of the model works. The below features—bedrooms, bathrooms, living area, and floors—are analyzed based on individual effect on pricing outcomes.

8.1 Bedrooms and Bathrooms: Quantity vs. Utility

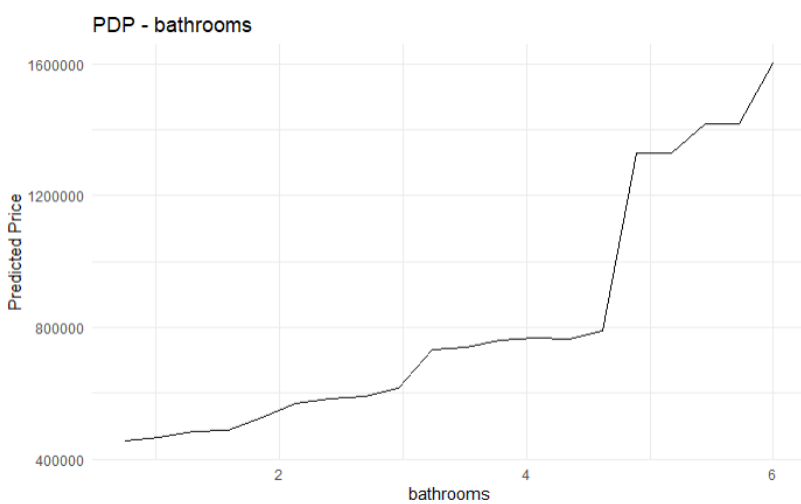
We begin by considering two of the most self-evident features of housing: bedrooms and bathrooms.

Bedrooms



The PDP for the number of bedrooms is positively correlated with the predicted price up to five bedrooms, where it levels off. This reflects declining returns: as more bedrooms undoubtedly contribute to increased valuations, the marginal value contribution of the extra bedroom declines after some point.

Bathrooms



In contrast, the PDP for bathrooms is more nonlinear and has a stronger impact. The model predicts a sharp increase in price with over four bathrooms, maybe indicating an entry into the luxury segment. This illustrates the model's ability to capture thresholds that correspond to real market expectations.

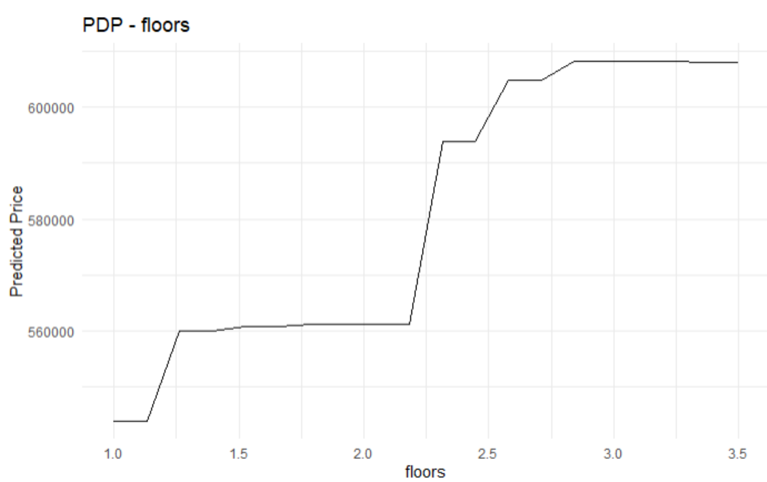
8.2 Living Space and Vertical Layout

This section deals with properties relating to size and structural layout.

Living Area (sqft_living)

The living area PDP shows a nearly linear predicted price increase with square footage up to around 7,500 sqft, and the effect levels off after that, indicating that large properties have less additional price per square foot considerably. This leveling off is common in real estate markets.

Number of Floors



The number of floors is a more stepwise effect on price. The biggest jumps appear between 1 and 1.5 floors, and again between 2 and 2.5 floors. It is possible that these steps reflect architectural class or zoning differences, and the additional value from floor counts seems to end at 3 floors.

9. Conclusion

The Partial Dependence Plots produced for the house price prediction model characterized the way structural features influence pricing. The number of bedrooms and bathrooms predictably affect price; they behave differently: bedrooms do experience diminishing returns after a certain number, while bathrooms have definite threshold effects related to luxury classification. Living area both has a strong effect and is nearly linear, with the highest space saturating; the number of levels creates a stepwise pricing behavior when moving from a one level home to more levels.

These results indicate the Random Forest model is effectively capturing differentiated market dynamics that include non-linear effects and feature interactions. PDP analysis supports interpretability of the model and confirms that the logic behind model predictions align with normal understandings of real estate value and the orientation of market features.