XAI: Model-agnostic methods

Exercise 5.- Model-agnostic: Partial Dependency Plot (PDP).

Remember: this exercise must have version-control using git and backup support through github. In the corresponding task, you only have to upload the link to the repository generated using git orders

1.- One dimensional Partial Dependence Plot.

The partial dependence plot shows the marginal effect of a feature on the predicted outcome of a previously fit model.

EXERCISE:

Apply PDP to the regression example of predicting bike rentals. Fit a random forest approximation for the prediction of bike rentals (**cnt**). Use the partial dependence plot to visualize the relationships the model learned. Use the slides shown in class as model.

QUESTION:

Analyse the influence of **days since 2011**, **temperature**, **humidity** and **wind speed** on the predicted bike counts.

2.- Bidimensional Partial Dependency Plot.

EXERCISE:

Generate a 2D Partial Dependency Plot with humidity and temperature to predict the number of bikes rented depending of those parameters.

BE CAREFUL: due to the size, extract a set of random samples from the BBDD before generating the the data for the Partial Dependency Plot.

Show the density distribution of both input features with the 2D plot as shown in the class slides.

TIP: Use geom_tile() to generate the 2D plot. Set width and height to avoid holes.

QUESTION:

Interpret the results.

3.- PDP to explain the price of a house.

EXERCISE:

Apply the previous concepts to predict the **price** of a house from the database **kc_house_data.csv**. In this case, use again a random forest approximation for the prediction based on the features **bedrooms**, **bathrooms**, **sqft_living**, **sqft_lot**, **floors** and **yr_built**.

Use the partial dependence plot to visualize the relationships the model learned.

BE CAREFUL: due to the size, extract a set of random samples from the BBDD before generating the data for the Partial Dependency Plot.

QUESTION:
Analyse the influence of **bedrooms, bathrooms, sqft_living** and **floors** on the predicted price.