Homework No. 4 - Ceteris Paribus, Partial Derpendence Plot

Paulina Kaczynska

November 27, 2022

The goal of this homework is to explain predictions of models trained on the Alcohol Effects on Study Dataset from Kaggle with Ceteris Paribus method and Partial Dependence method. It bases on the data preprocessed in previous homeworks.

1 Task 1

$$f(x_1, x_2) = (x_1 + x_2)^2 \tag{1}$$

Partial Dependence profile for x_1 is calculated in the following way:

$$g^1_{PD}(z) = E_{x_2} f(x_1 = z, x_2) = \int_{-1}^{1} \frac{1}{2}(z + x_2)^2 dx_2 = \frac{1}{2}(\frac{x_2^3}{3} + x_2^2 z + x_2 z^2)|_{-1}^{1} = \frac{1}{3} + z^2 \tag{2}$$

The Marginal Effects are calculated in this way:

$$g^1_{ME}(z) = E_{x_2|x_1=z} f(x_1 = z, x_2) = \int_{-1}^{1} \frac{1}{2}\delta(x_2 = z)(z + x_2)^2 dx_2 = 4z^2 \tag{3}$$

Accumulated local effects are calculated here:

$$g^1_{ALE}(z) = \int_{-1}^{z} E_{x_2|x_1=y} \frac{\partial f(x_1, x_2)}{\partial x_1} dy = \int_{-1}^{z} E_{x_2|x_1=y} 2(x_2 + x_1) dy = \int_{-1}^{z} 4y dt = 2y^2|_{-1}^{z} = 2z^2 - 2 \tag{4}$$

2 Task2

2.1 Ceteris Paribus Explanation for random forest regressor

Firstly, I train Random Forest Regression using function from sklearn package. This method is a tree ensemble method.

I sample four students and explain predictions with the Ceteris Paribus method with the use of DALEX package. Students' predicted grades are: 11.17, 6.07, 12.44 and their numbers are: 936, 148, 480 Their true grades are: 11, 0, 12.

The plots of what if explanation for the number of absences and studytime can be seen on the Figure 1. Number of absences contributes differently for students 936 and 148: for the first one the smaller number of absences, the bigger the grade and for the latter it is the opposite way.

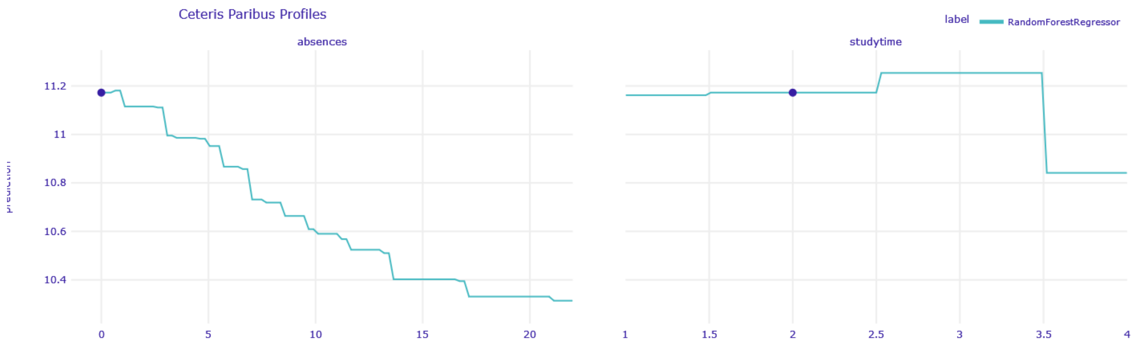
2.2 Partial Dependence explanation of random forest regressor

I calculated the global explanation of this model with Partial Dependence Plot. It can be seen on the Figure 2. It is very interesting that the number of absences equal to zero contributes very negatively to the overal score, while for every other number of absences the smaller the number, the better the score. The aggregated results show that unsurprisingly the partial dependence plot has very small range of values comparing to individual profiles of ceteris paribus. Partial dependence plot for studytime variable shows how the only difference that matters is whether reported studytime is bigger than 2.

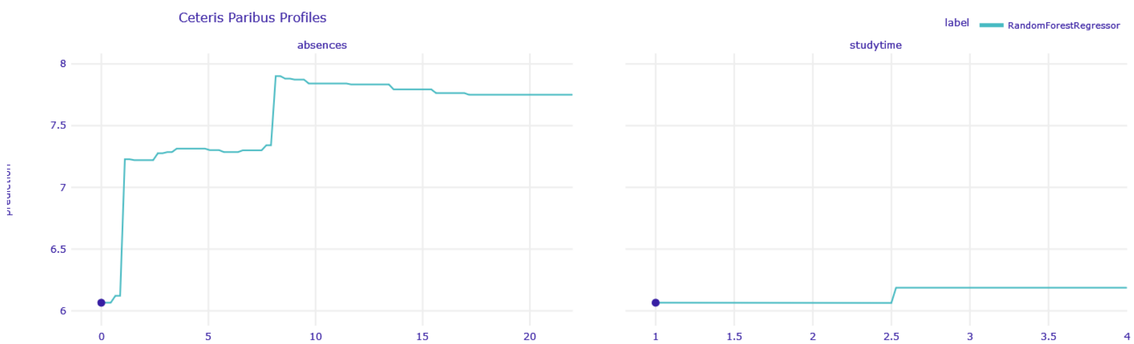
2.3 Partial Dependence for linear regression and stochastic gradient descent

The Linear Regression model and Stochastic Gradient Descent model on the same data were trained and explained with Partial Dependence Plot from DALEX package. The results can be seen on the figure 3.

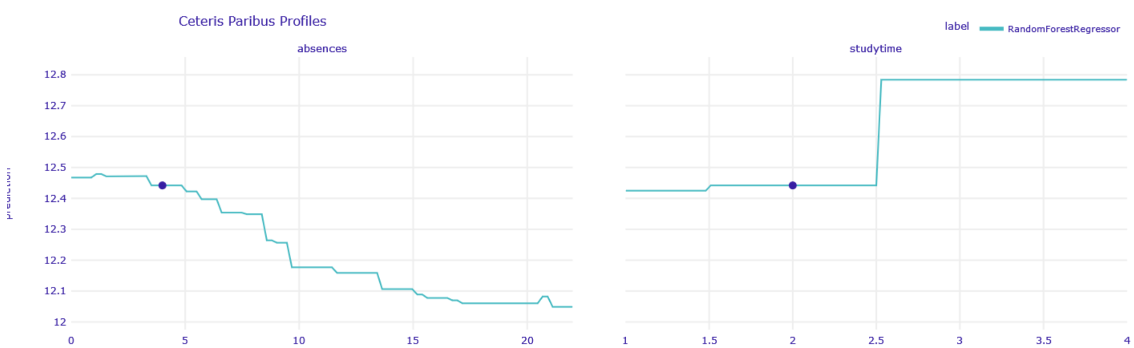
The dependence of the final score on the feature is linear in both cases. It is obvious, since the architectures underlying these models assume linear dependence on the variables. SGD plots for absences and studytime are steeper, so the coefficients' moduls are larger.



(a) Student No. 936



(b) Student No. 148



(c) Student No. 480

Figure 1: A prediction for three random students made with Random Forest Regression explained with Ceteris Paribus

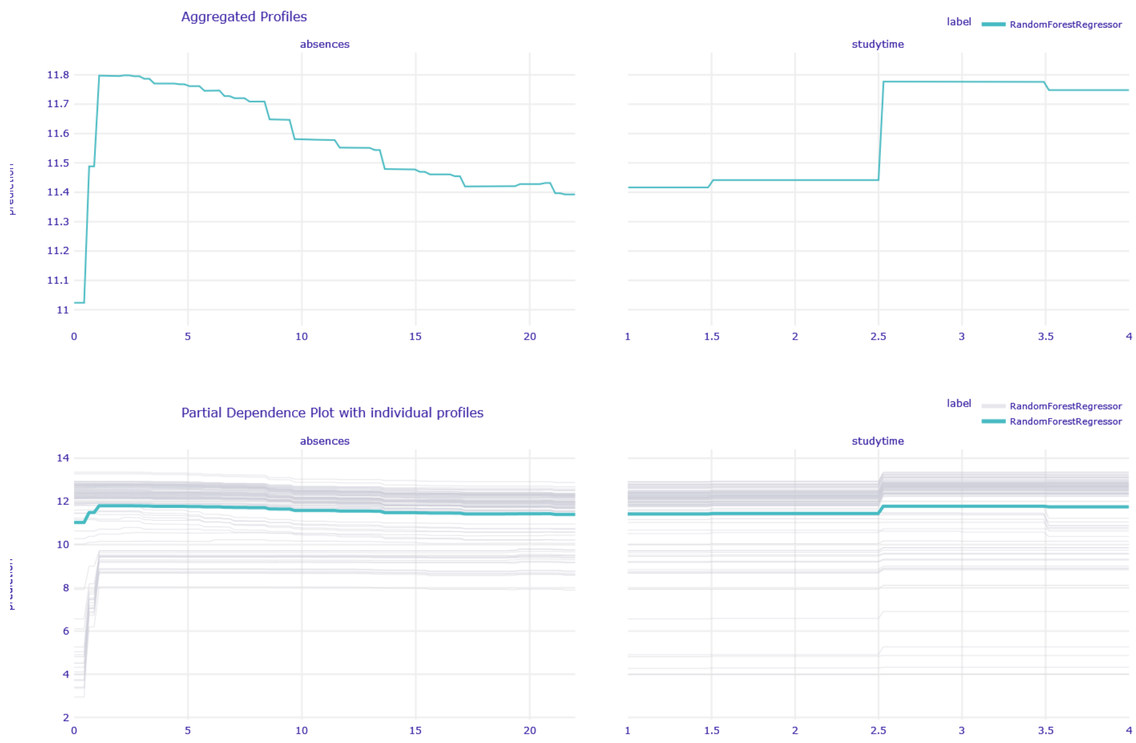
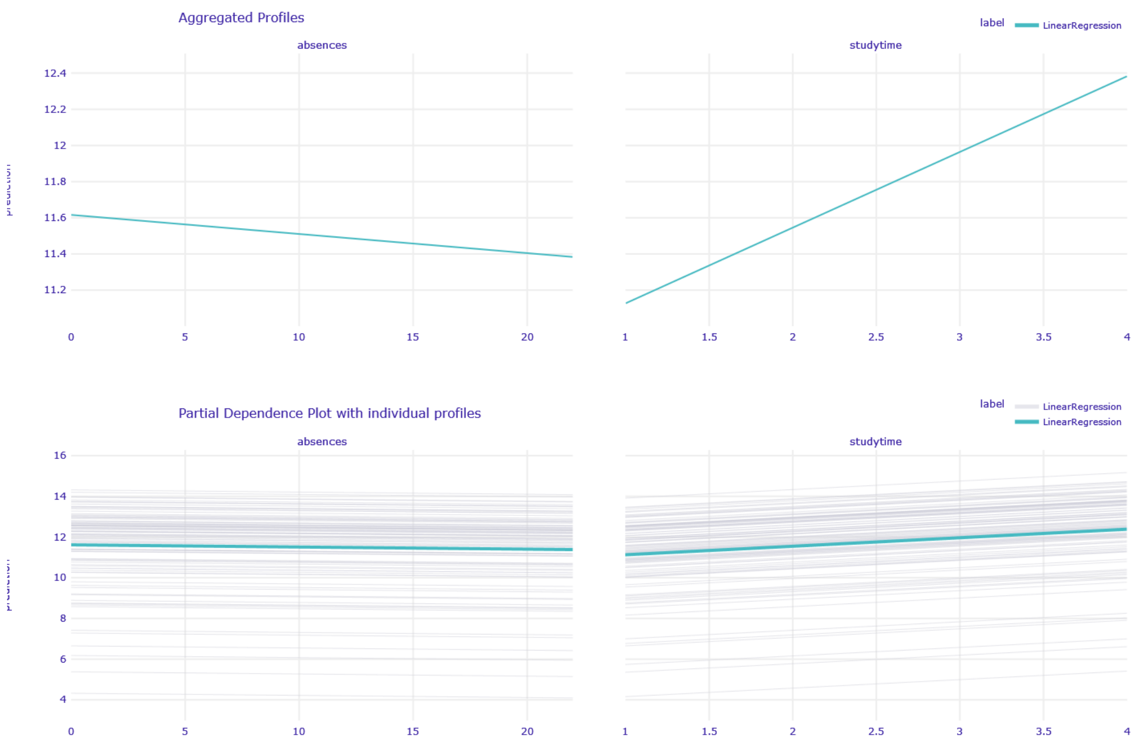


Figure 2: Partial Dependence plot of random forest regressor



(a) Partial dependence plots for Linear Regression

(b) Partial dependence plots for Stochastic Gradient Descent

Figure 3: Partial Dependence explanations of the same features made for linear regression and stochastic gradient descent.