

Homework No. 3 - LIME

Paulina Kaczynska

October 2022

The goal of this homework is to explain predictions of models trained on the Alcohol Effects on Study Dataset from Kaggle with LIME method and compare it with SHAP method. It bases on the data preprocessed in previous homeworks.

1 LIME on Random Forest Regression

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 LIME method with the use of DALEX package. Students' predicted grades are: 12.42, 14.73, 3.2, 12.31 and their numbers are: 470, 198, 239, 73. Their true grades are: 11, 18, 0, 14.

The plots of contribution of the features to the prediction can be seen on the Figure 1.

In all cases the number of failures is the most important feature. Absences and receiving school support are also on high positions. However, because of the randomness, the exact values for the same student are not the same when explanation is repeated, as can be seen for the first student in the sample: the failures' contribution firstly was 3.95 and after repeating it it was 3.81.

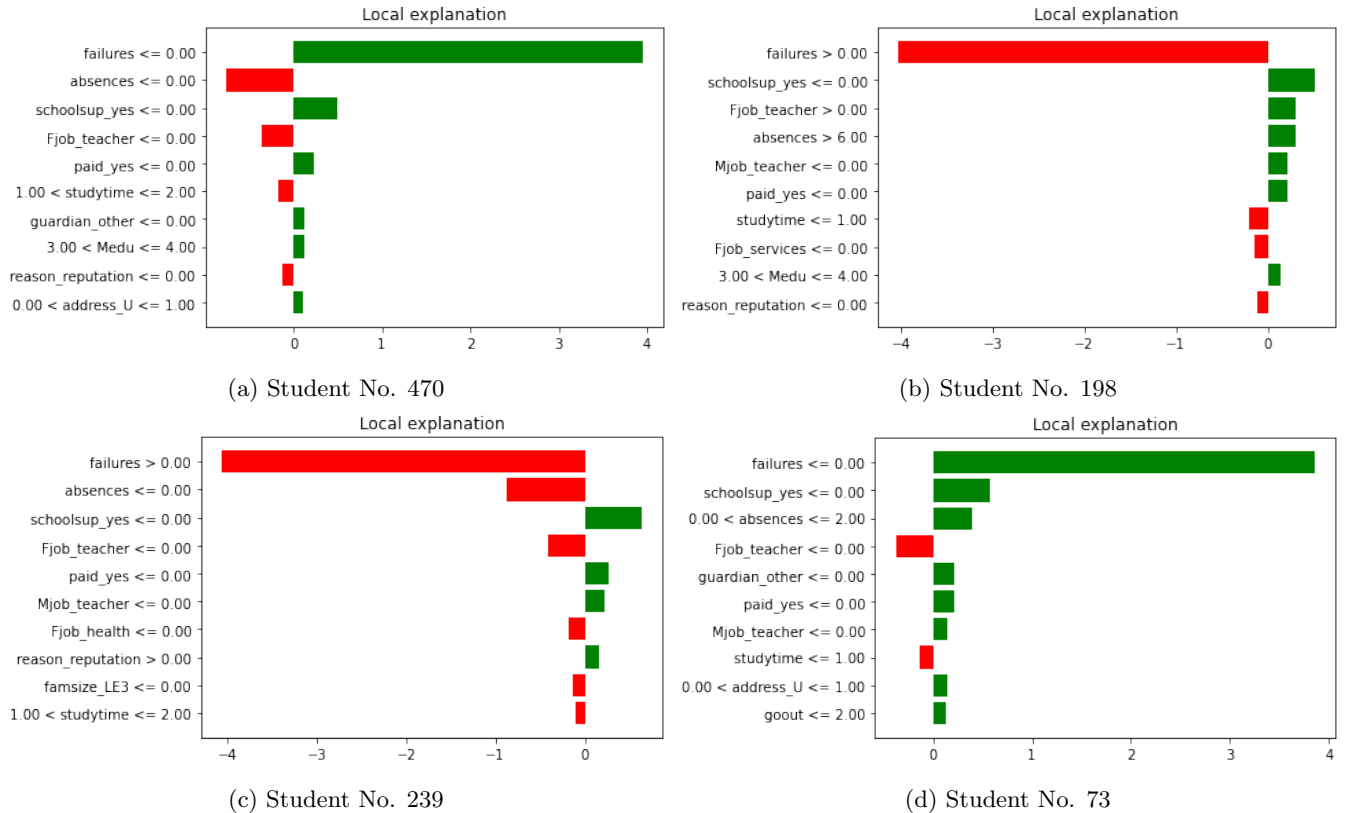


Figure 1: A prediction for four random students made with Random Forest Regression explained with LIME

2 Comparison of LIME explanation with SHAP explanation

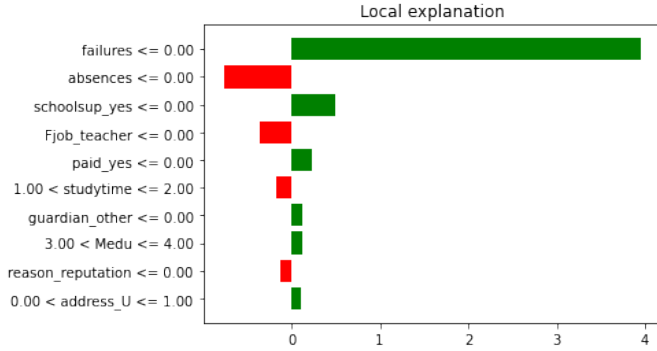
I calculate the explanation of the same model with SHAP method using shap package. The comparison of this results with the LIME explanation can be seen on the Figure 2.

The explanations differ in SHAP and LIME. For example student no. 198 has the different feature with the biggest contribution to the explanation in both methods: failures in LIME and father working as a teacher in SHAP. I would propose a hypothesis, that lime is more often proposing the feature that is most important in other observations because of the sampling needed to construct simplified model.

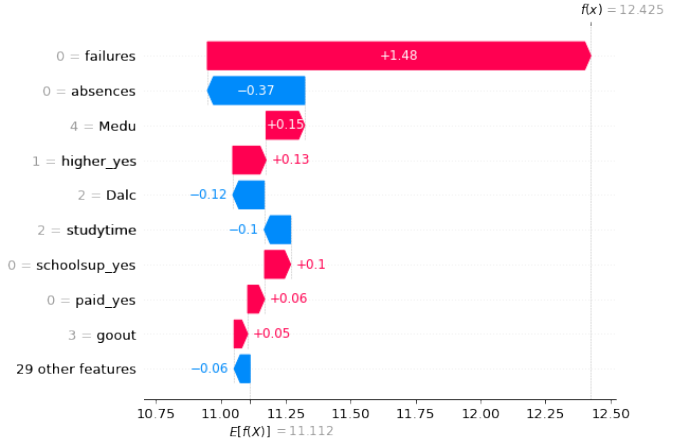
3 Comparison of LIME explanation on different models

The Linear Regression model on the same data was trained and explained with LIME method from DALEX package. The results can be seen on the figure 3.

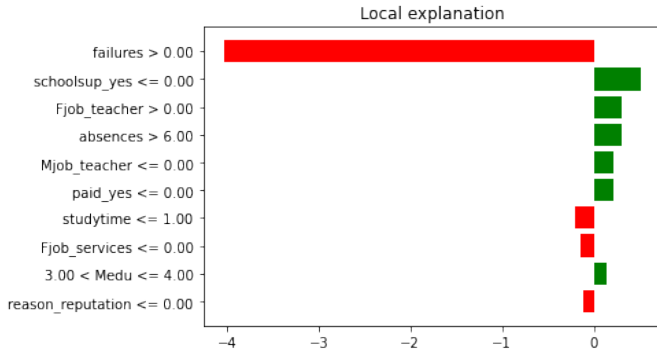
The values for the features other than the most important are higher for the linear regression than the random forest regression. It was similar in the previous homework when comparing explanation of these models made with the SHAP method. Explanations of linear regression also favour features more important in other observations as was seen previously.



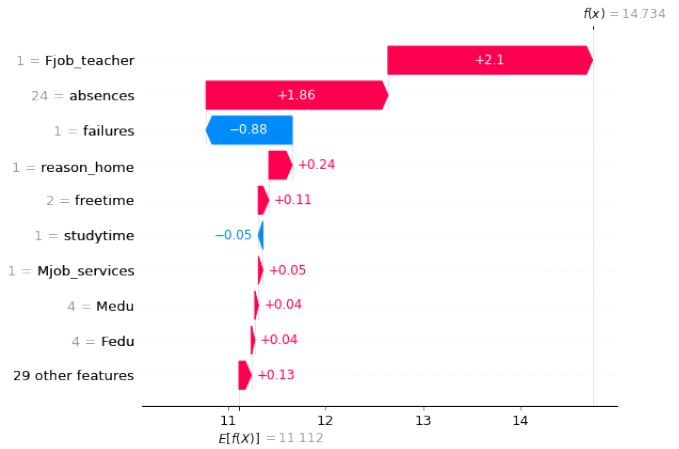
(a) Student No. 470 - LIME



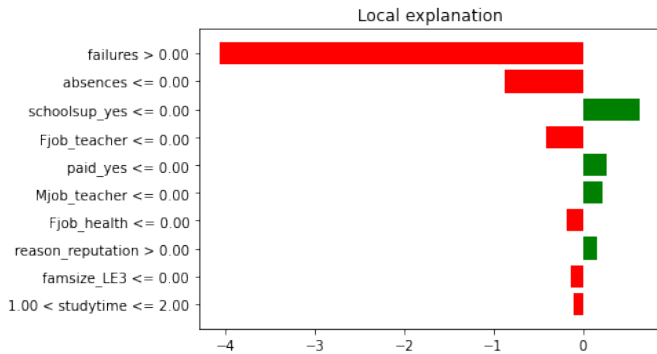
(b) Student No. 470 - SHAP



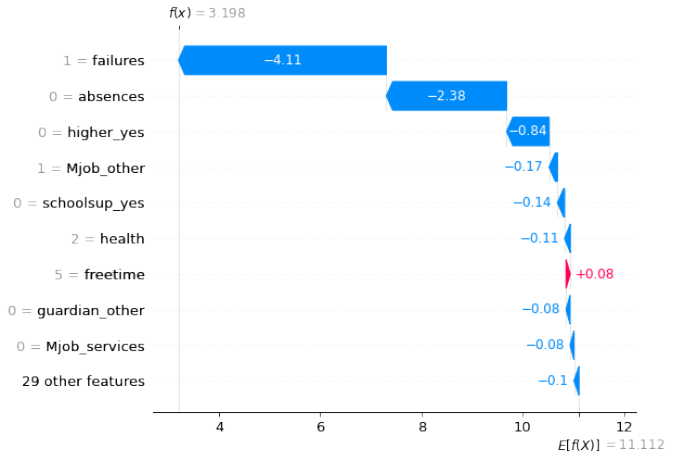
(c) Student No. 198 - LIME



(d) Student No. 198 - SHAP

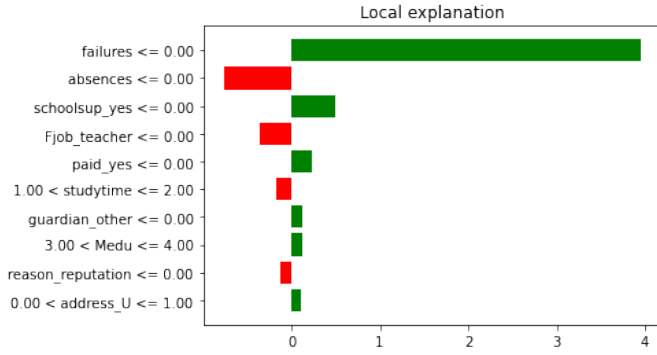


(e) Student No. 239 - LIME

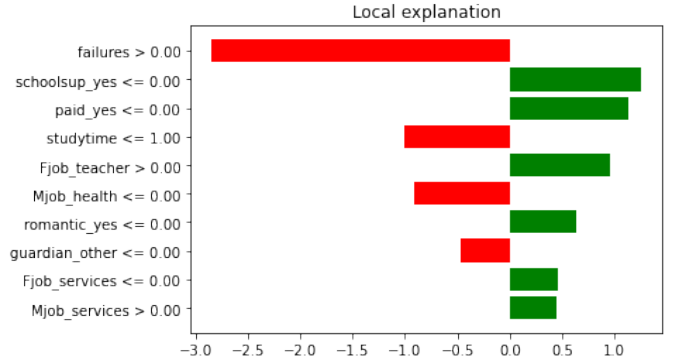


(f) Student No. 239 - SHAP

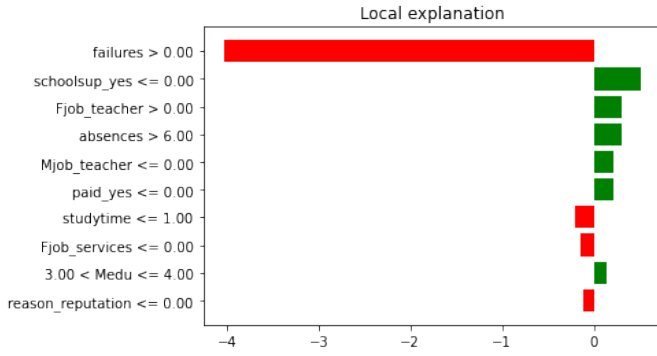
Figure 2: A comparison of the explanations with LIME and with SHAP made on the same Random Forest Regression model



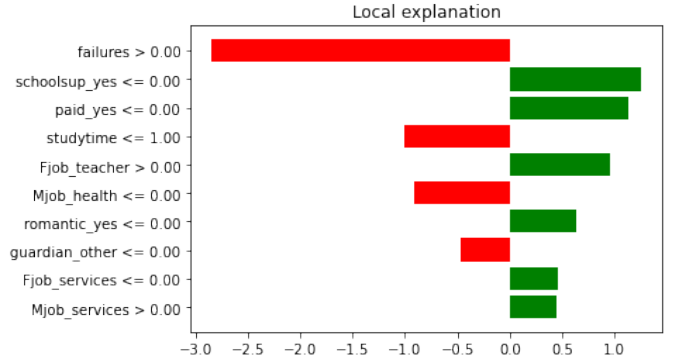
(a) Student No. 470 - Random Forrest Regression



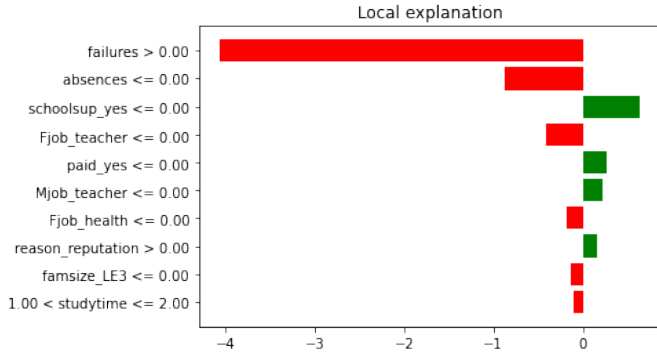
(b) Student No. 470 - Linear Regression



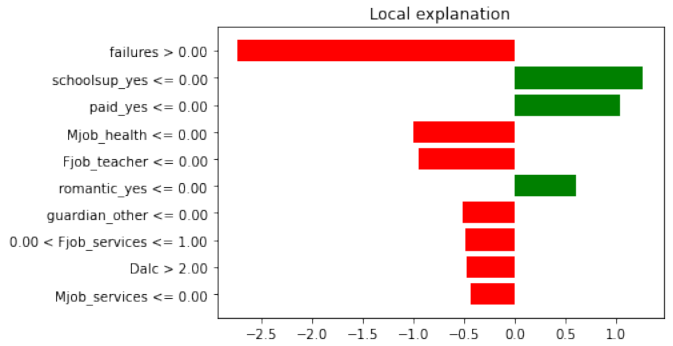
(c) Student No. 198 - Random Forrest Regression



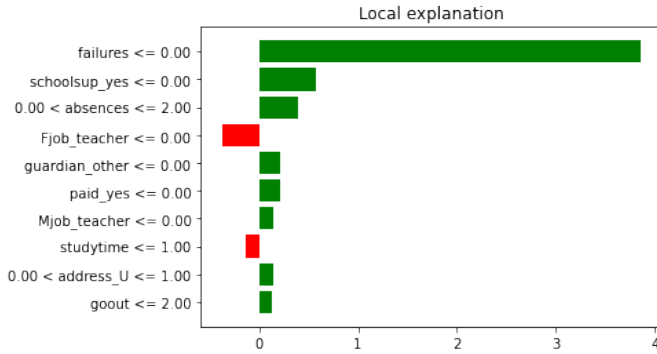
(d) Student No. 198 - Linear Regression



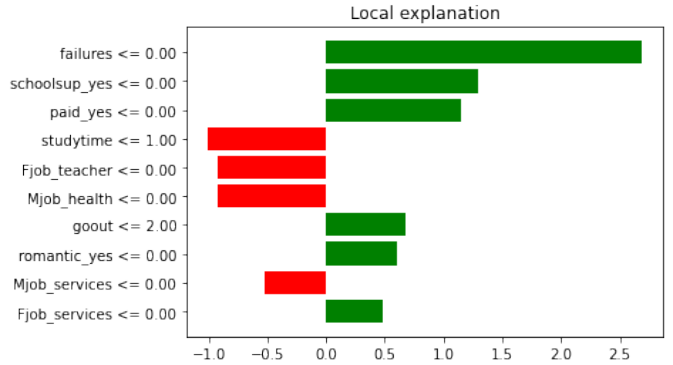
(e) Student No. 239 - Random Forrest Regression



(f) Student No. 239 - Linear Regression



(g) Student No. 73 - Random Forrest Regression



(h) Student No. 73 - Linear Regression

Figure 3: A comparison of the explanations with LIME made on the two models: Random Forest Regression and Linear Regression