Explainable AI on machine learning model

Mateusz Cedro

Goal

The goal is to find out the way in which insurance companies charge their clients with regard to their social status and health conditions.

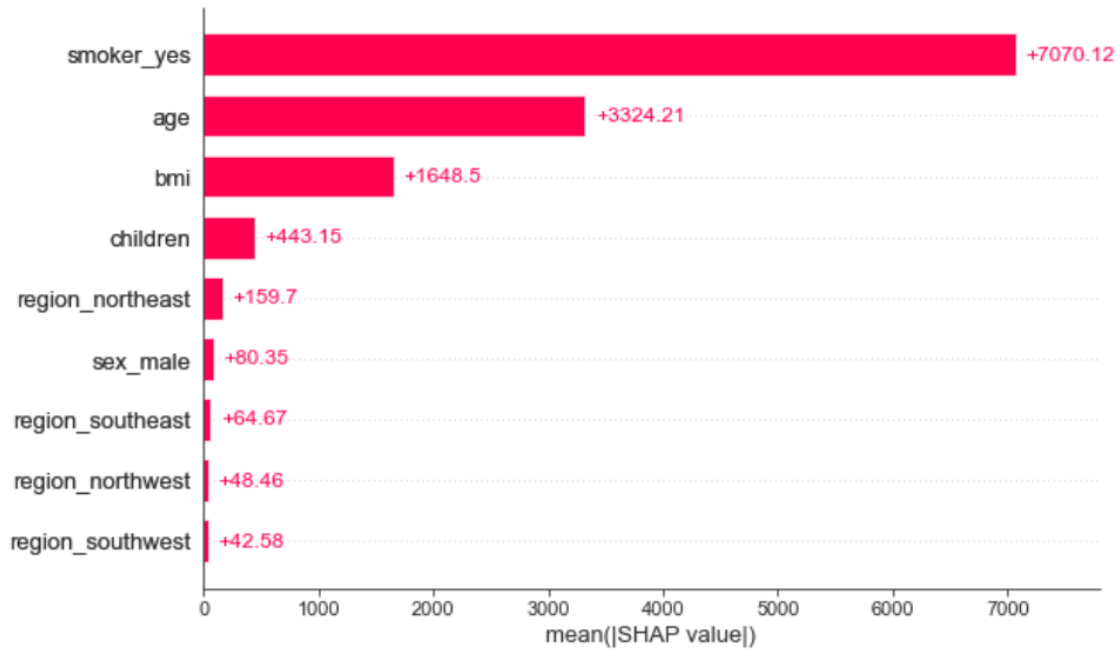
Dataset

The US Health Insurance Dataset contains data on 1338 people with regard to Age**,** Sex**,** BMI, number of Children, if the person is a Smoker, and Region of living. Charges were the target variable which indicates the cost of insurance levy.

Explainable Artificial Intelligence (XAI)

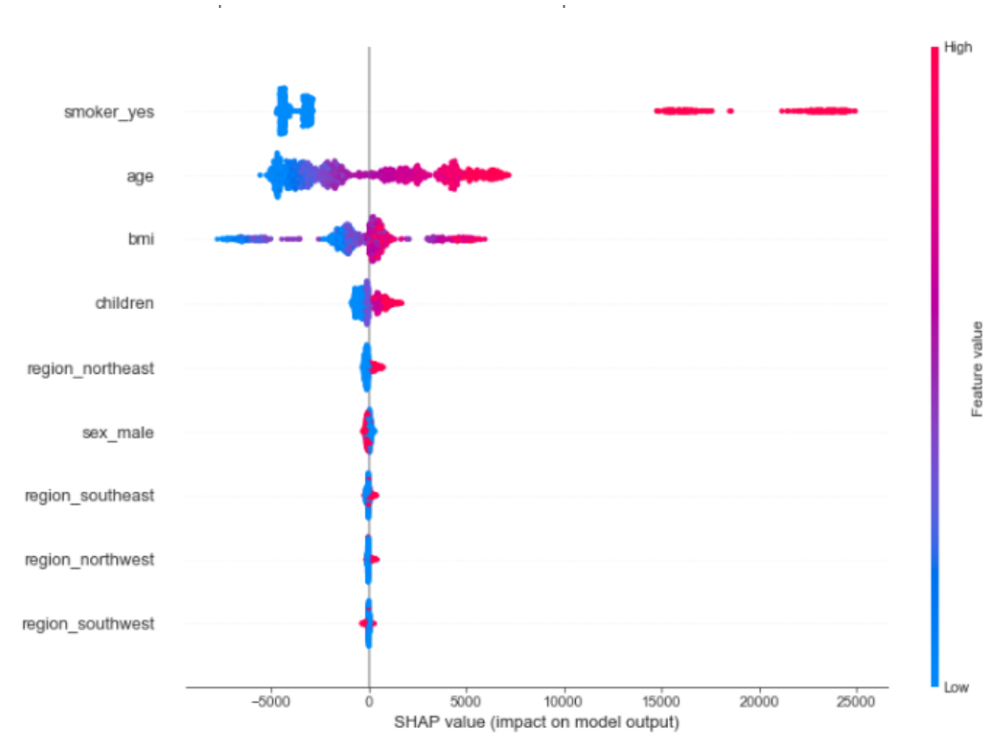
For the best performing model, which was a Random Forest regression model, I have done the XAI model explanations with SHAP library which helps in model explanations basing on Shapley values and feature attribution.

The plot of variable importance is presented below:



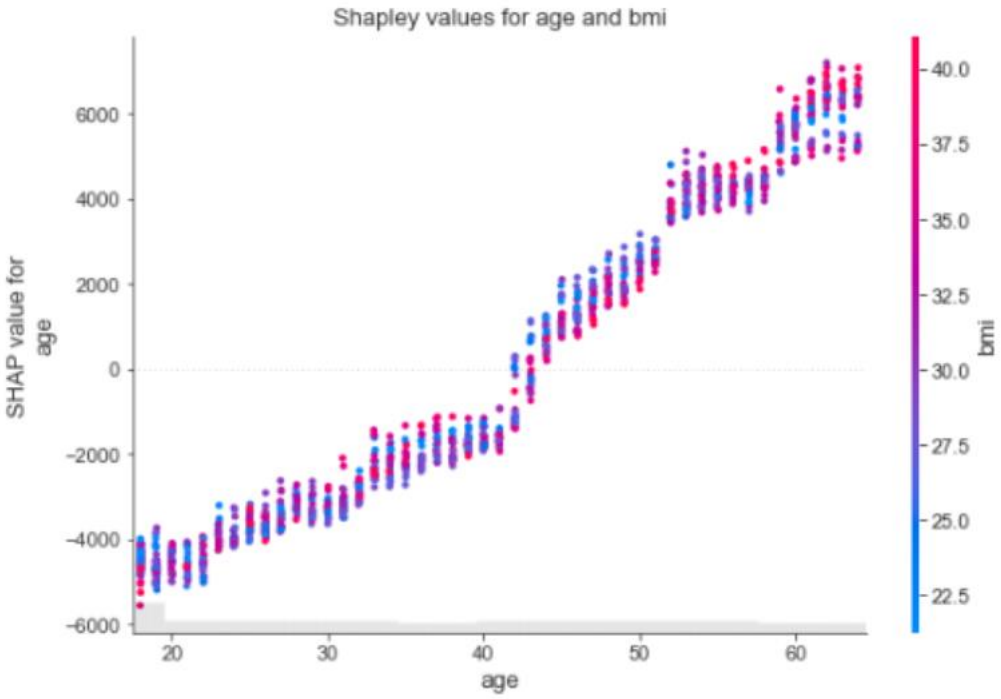
We can clearly see that the variables which contributes most in calculating mean Shapley values for the Charges levy are Somker_yes, AGE, and BMI, where the smoker contributes twice as much as AGE, and fourth as much as BMI. Interestingly, Sex variable contributes less than region of Northeast of living.

Below the more detailed plot with distributions of variables are presented:



We can clearly see the distinction for Shapley values for each variable

Below we can see the plot describing the AGE and BMI Shapley values:

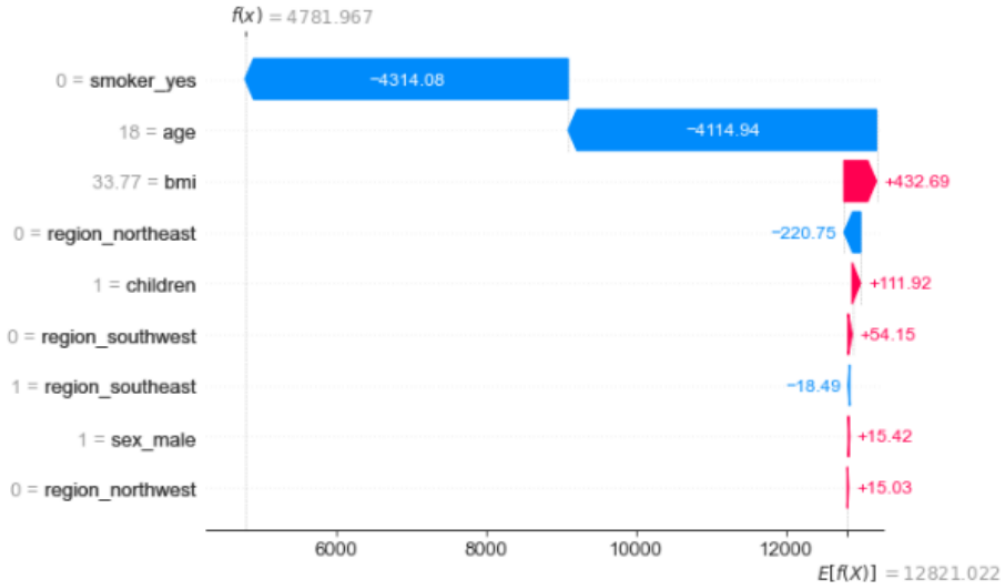


We can clearly see that the age is strictly positively correlated with the Shapley values for additional value of Charges levies. For Age=43 we can see that Shapley values are around

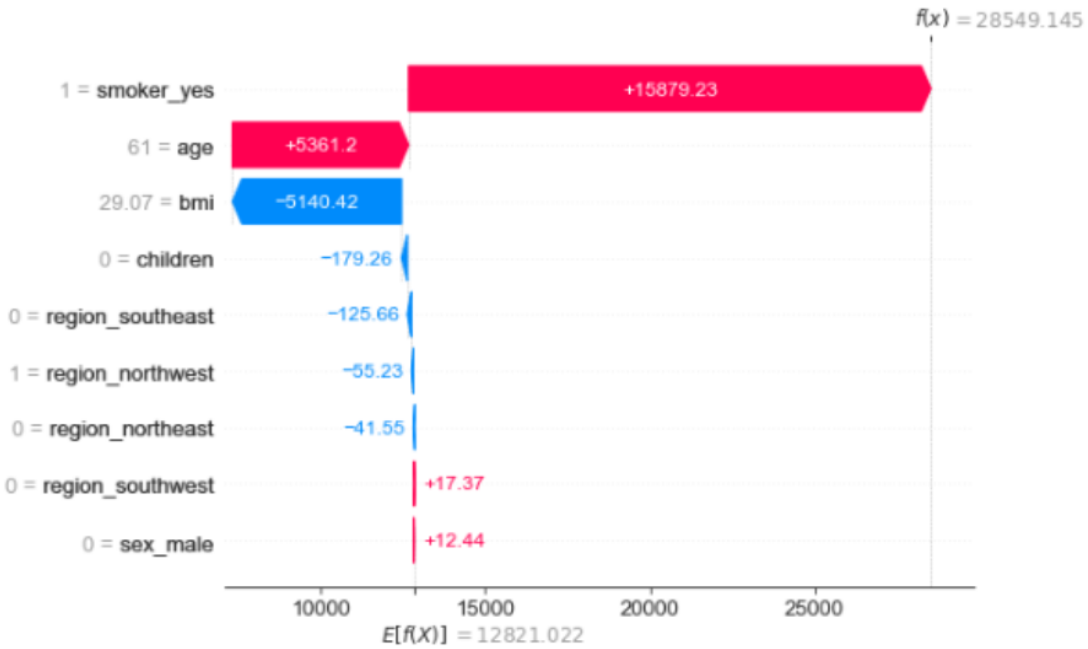
zero; age below 43 decreases the insurance charges amount, and higher age than 43 increases the values for insurance levies.

Finally, we would like to present two observations:

- First: 18 years old non-smoker



- Second: 61 years old smoker



We can see the exact differences between these two observations

Conclusions

In this report I have presented the approach to explain complex artificial intelligence model with the use of Shapley values, having previously train and tune hyper-parameter model.

The best performing model was Random Forest model which performed better than XGboost and Multi-Layer Perceptron models.