Funnel Plot: How to compare predictive models?

Szymon Maksymiuk[1], Przemysław Biecek[1,2]

[1]Faculty of Mathematics and Information Science, Warsaw University of Technology

[2]Faculty of Mathematics, Informatics, and Mechanics, University of Warsaw

DALEXtra

DALEXtra is a R package that provides tools to compare and interprate various machine learning models. Nowadays, there is a strong belief that we need to strike a balance between interpretability and accuracy. However, in field of the interpretable machine learning, there are more and more new ideas for explaining black-box models, that are implemented in R, Python or Java. DALEXtra let us import all those models into R environment in order explain or comapre in a unified way. Package also provides new approach to comparison of models' scores.

Funnel Plot

Model agnostic comparison of models is hard task and many reaserchers seek for a solution for that issue. It is becasue we would like to somehow compare complaetly different methods like neural networks and random forests. Plain accuracy measure, for instance AUC or RMSE my not be the best solution as they treat model globally and thus may mislead us.

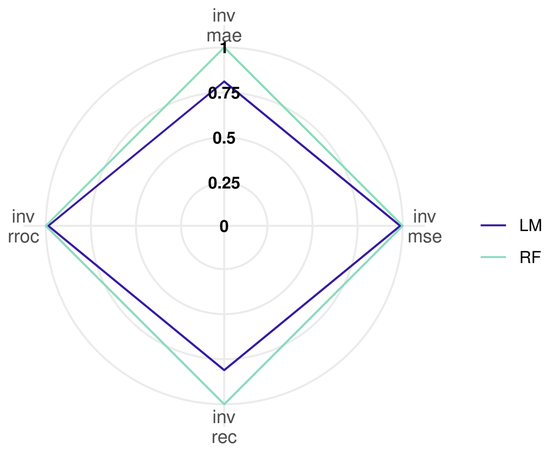

Figure 1: In above radar plot we can see measures of two different models RF - which is Random Forest and LM which is Linear Model. As we can see Mean Square Error for both of them is nearly the same what may state they are equaly good

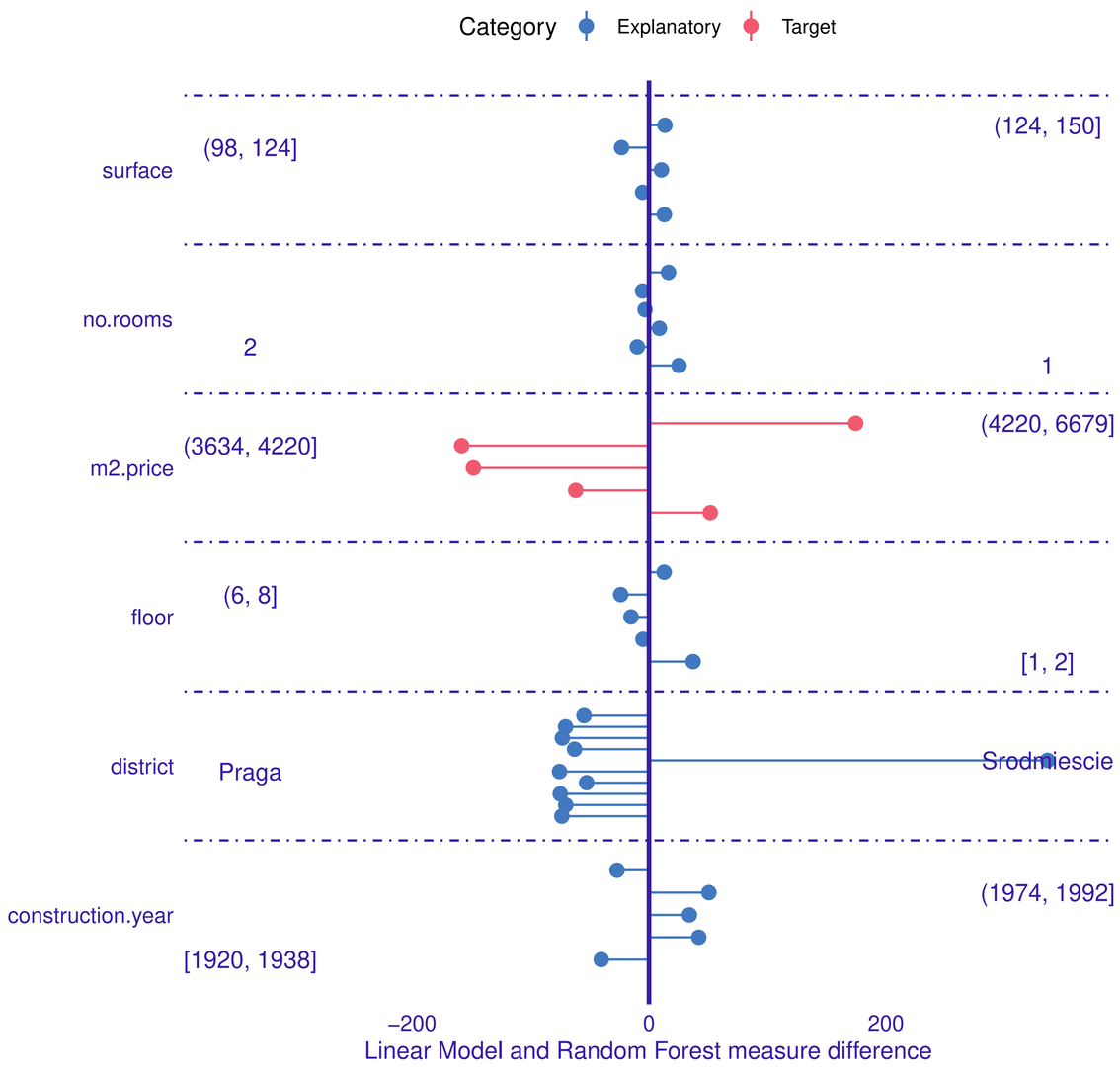

Figure 2: Funnel Plot. For every variable, dot on the right side of the violet line means that Linear Model is better for given subset. Dot on the left means that Random Forest model is better.

Above Funnel Plot help us to understand how model predictions behave for various subsets of test data. It allows us to suit model for our exact purposes, for the instance we would definitely use Linear Model if we would like to predict apartment located in Srodmiescie. That easy to understand conception extracts from prediction scores information that would never be accessible using plain preformance like in the 1st figure.

Champion Challenger

Champion Challenger analysis is a process, that helps us find the best model by comparing a current best solution (Champion) with other models (Challengers). The core of it is a Funnel Plot accompaniated by overall analysis of measurements (Fig. 1), response accoradance analysis (Fig. 3) and training-test coamprison that that allow us to overcome overfitting. DALEXtra helps with whole analysis by creating automated report that can be seen by scanning below QR code.

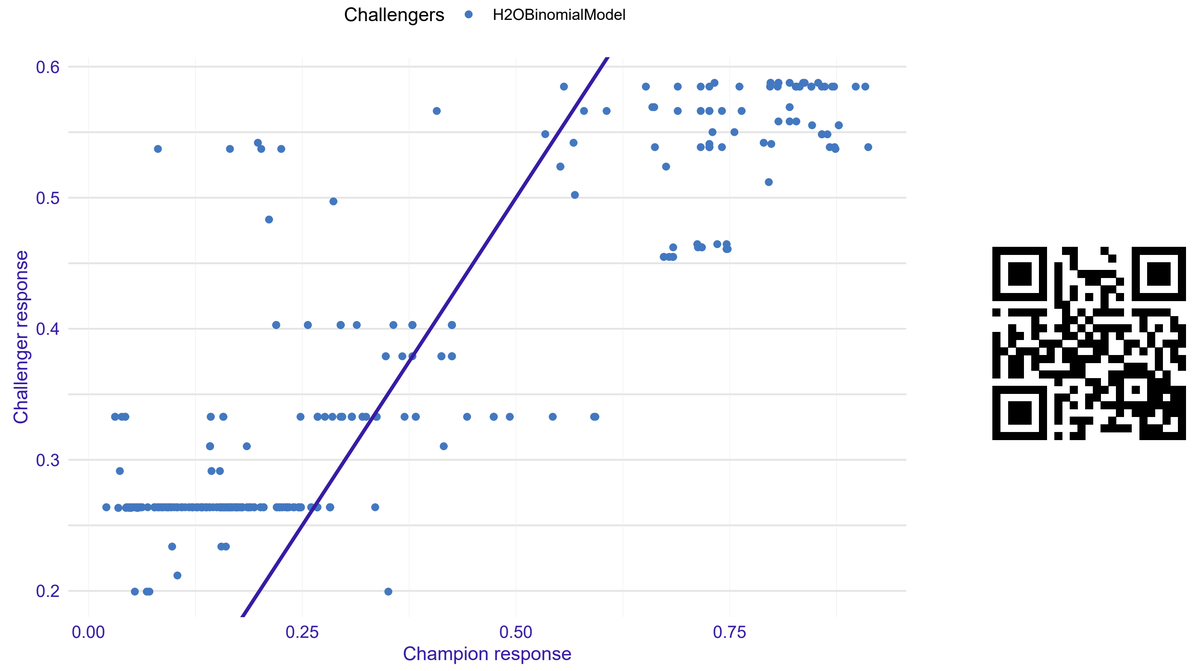

Figure 3: In this example Champion model which responses are present on the OX is scikit-learn GBM while Challenger is h2o GBM. That simple plot allow as to detect bad behaviour of models. Here we have huge diversity of scikit-learn responses while h2o's set of unique responses is quite thight. This suggests that something may be wrong and acctually is, since h2o was fitted with only 500 trees while scikit-learn with 5000

Integration

All of anlysis presented here, can be accessed despite of origin of our model. DALEXtra integrates potentially different environmetns and allows us to establish common ground between them to compare models built in Python, Java h2o or even MLJAR in a unified way, without platform bias. Claims that your model is better only becasue was build in H2O automl are no more!

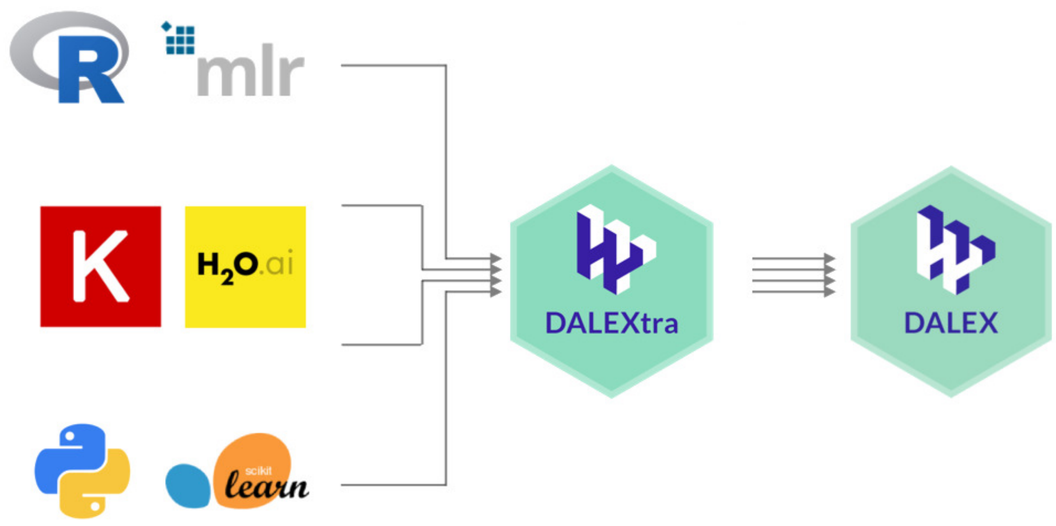

Figure 4: Example of workflow with DALEXtra. We can use many different environments and create DALEX explainer

References

[1] Przemyslaw Biecek. Dalex: Explainers for complex predictive models in r. *Journal of Machine Learning Research*, 19(84):1–5, 2018. URL http://jmlr.org/papers/v19/18-416.html.

[2] Alicja Gosiewska and Przemysław Biecek. auditor: an R Package for Model-Agnostic Visual Validation and Diagnostics. *The R Journal*, 2019. doi: 10.32614/RJ-2019-036. URL https://journal.r-project.org/archive/2019/RJ-2019-036/index.html.

[3] Szymon Maksymiuk, Przemysław Biecek, and Katarzyna Pękała. DALEXtra: an extension for DALEX package. *CRAN*. URL https://cran.r-project.org/package=DALEXtra.

Acknowledgements

Work on this package was financially supported by the 'NCN Opus grant 2016/21/B/ST6/02176'.

▶ ✉ sz.maksymiuk@gmail.com