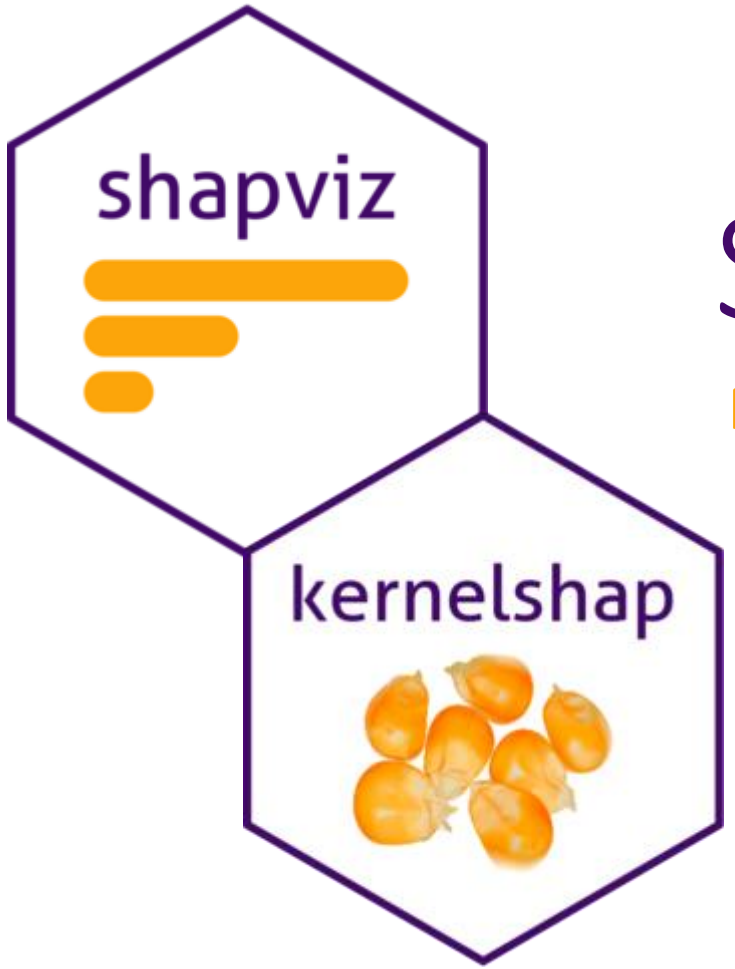


SHAP in R

Michael Mayer

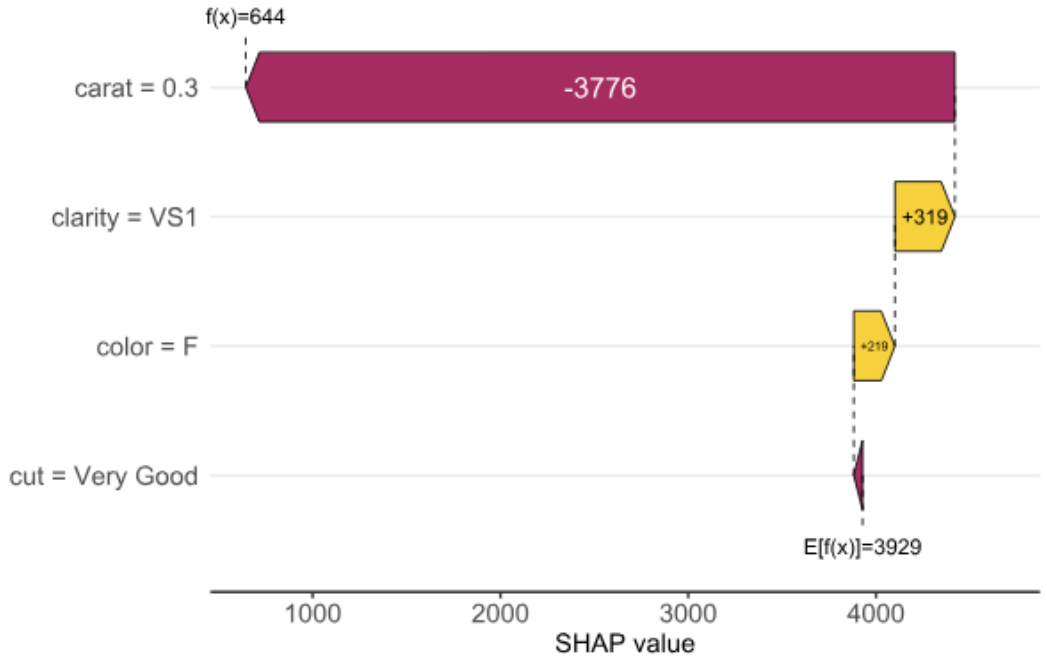
SHAP Everywhere

- SHapley Additive exPlanations (Lundberg & Lee, 2017)
- Fantastic Python library

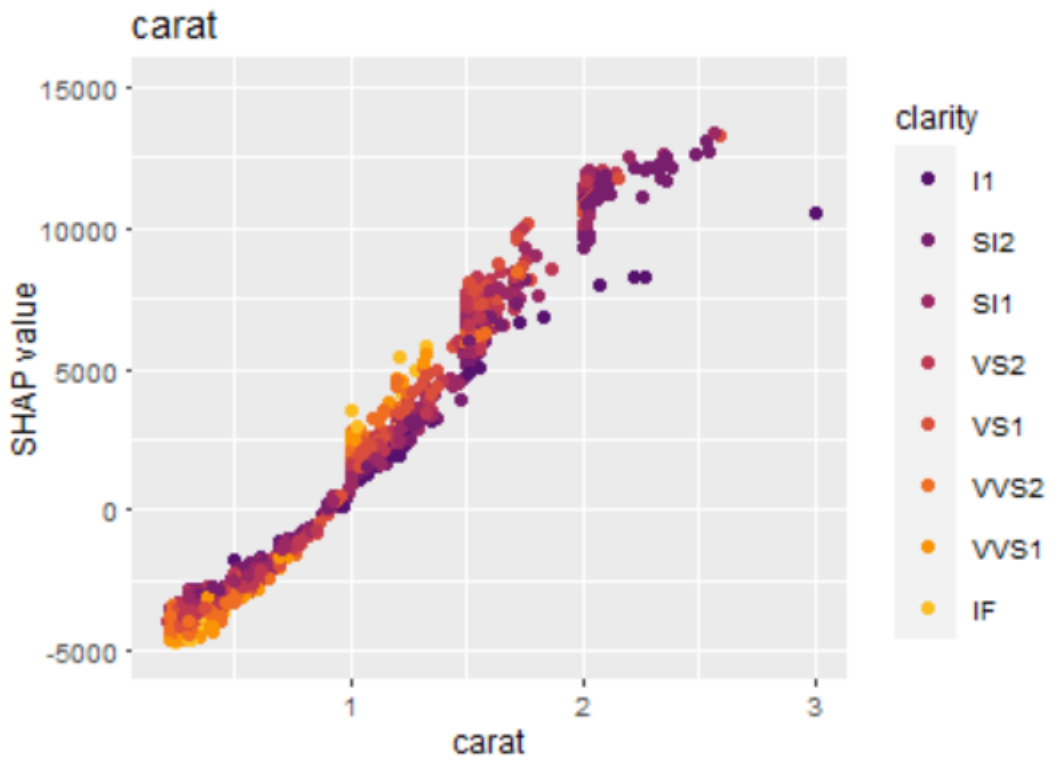
R is catching up

SHAP 1x1

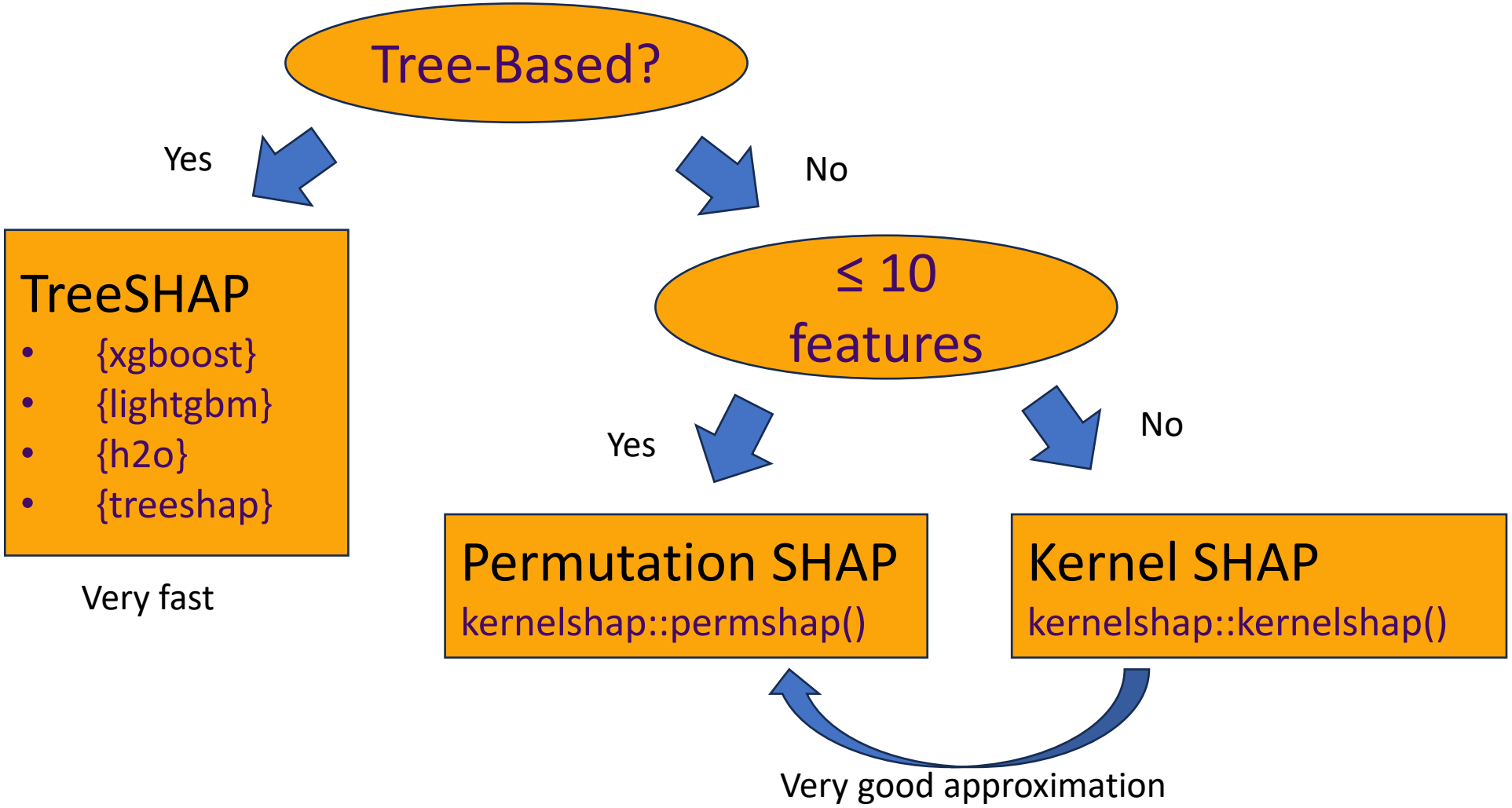
SHAP decomposes a model prediction fairly into additive contributions of the features



Repeat for many predictions to describe model as a whole



Three Algorithms to Crunch SHAP Values



~~One Ring to Rule them all~~

One Package to Plot them all: {shapviz}

treeshap

LightGBM

H2O.ai

dmlc XGBoost

shapr

{shapviz} API

data.frame-like with feature values

1. Initialize «shapviz» object

```
shap_values <- shapviz(object, X, …)
```

- Matrix of SHAP values
- Model: XGBoost/LightGBM/h2o-Booster, needs also `X_pred`
- Output of {kernelshap}, {fastshap}, {treeshap}, {shapr}, …

2. Use plot functions (ggplot2)

- `sv_importance(shap_values, …)`
- `sv_dependence(shap_values, v = «carat»)`
- …

Demo in R

Things to Explore

- SHAP interactions
- Multivariate output (e.g., multi-class classification)
- How do TreeSHAP, Permutation SHAP, Kernel SHAP work?
- Feature construction is relevant: not too correlated features!

A SHAP analysis is as good or bad as your model!

Questions?

Resources

- Slides and demo: https://github.com/mayer79/demo_shapviz

- {shapviz}: https://github.com/ModelOriented/shapviz
- {kernelshap}: https://github.com/ModelOriented/kernelshap
- Python: https://github.com/shap/shap
- Theory and code in R and Python:
  SHAP for Actuaries: Explain any Model (2023)
  Michael Mayer, Daniel Meier, and Mario V. Wüthrich
  http://dx.doi.org/10.2139/ssrn.4389797
- Original SHAP paper:
  A unified approach to interpreting model predictions (2017)
  Scott Lundberg and Su-In Lee
  NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems