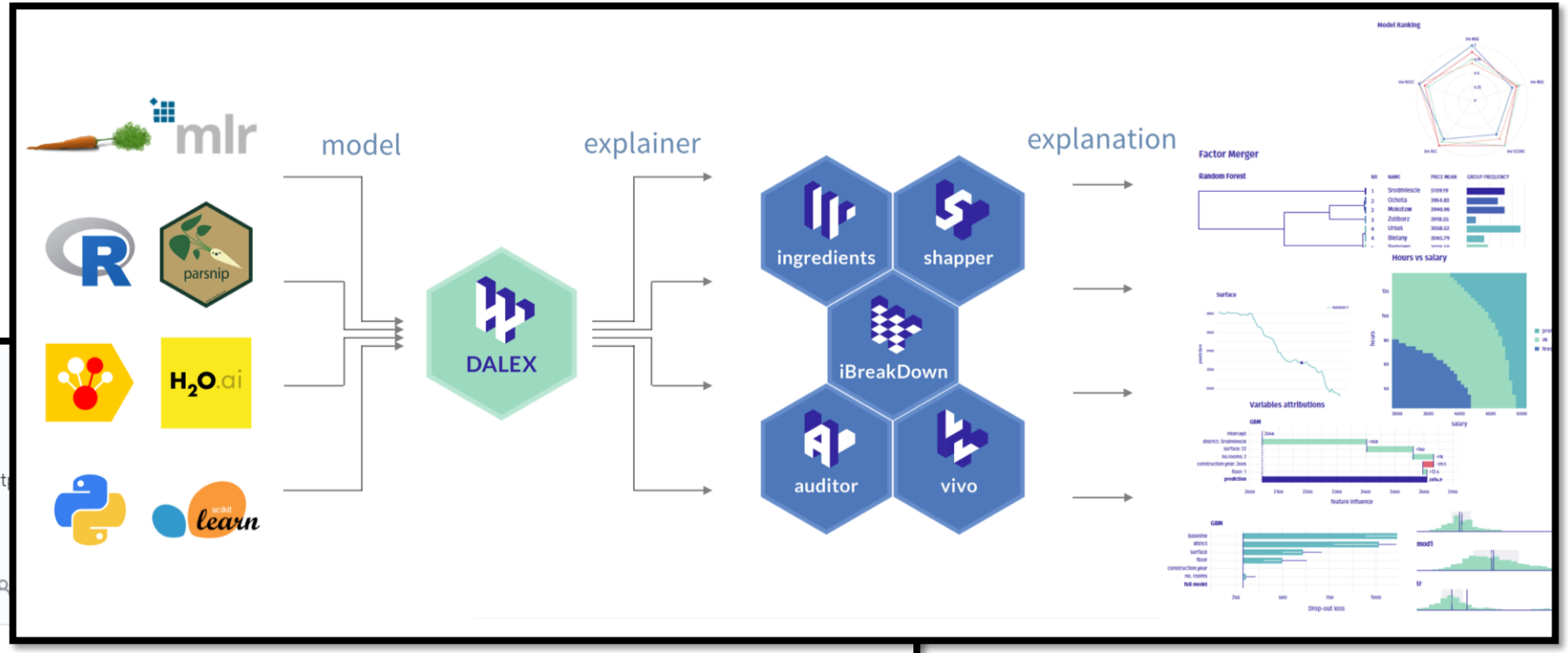
What’s new in DrWhy.AI?

Hubert Baniecki, Warsaw, 09.2020

DrWhy.AI



Model Oriented

MI2DataLab @ Warsaw University of T... http

Repositories 45 Packages People 21

Pinned repositories

DALEX
moDel Agnostic Language for Exploration and eXplanation
Python 649 97

DrWhy
DrWhy is the collection of tools for eXplainable AI (XAI). It's based on shared principles and simple grammar for exploration, explanation and visualisation of predictive models.
R 381 51

randomForestExplainer
A set of tools to understand what is happening inside a Random Forest
R 162 25

modelStudio
Interactive Studio for Explanatory Model Analysis
R 134 16

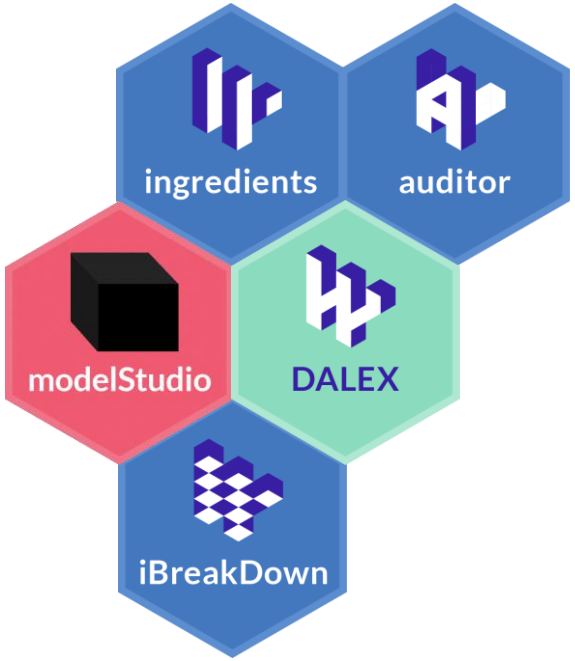
modelDown
modelDown generates a website with HTML summaries for predictive models
R 99 11

iBreakDown
Break Down with interactions for local explanations (SHAP, BreakDown, iBreakDown)
R 53 8

http://drwhy.ai/

Hi!

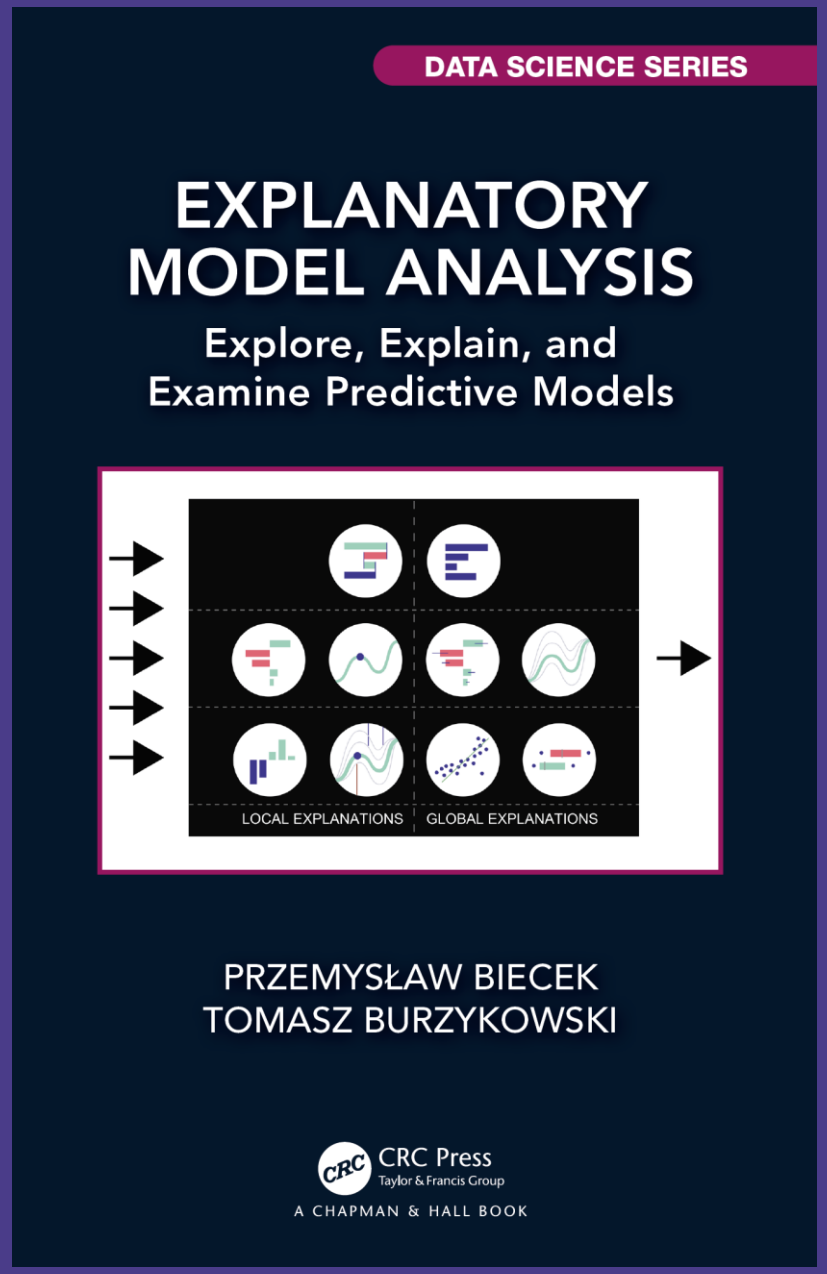
- Data Science student at Warsaw University of Technology
- contributing and maintaining the DrWhy.AI universe
- interested in Explainable AI and model-human interaction



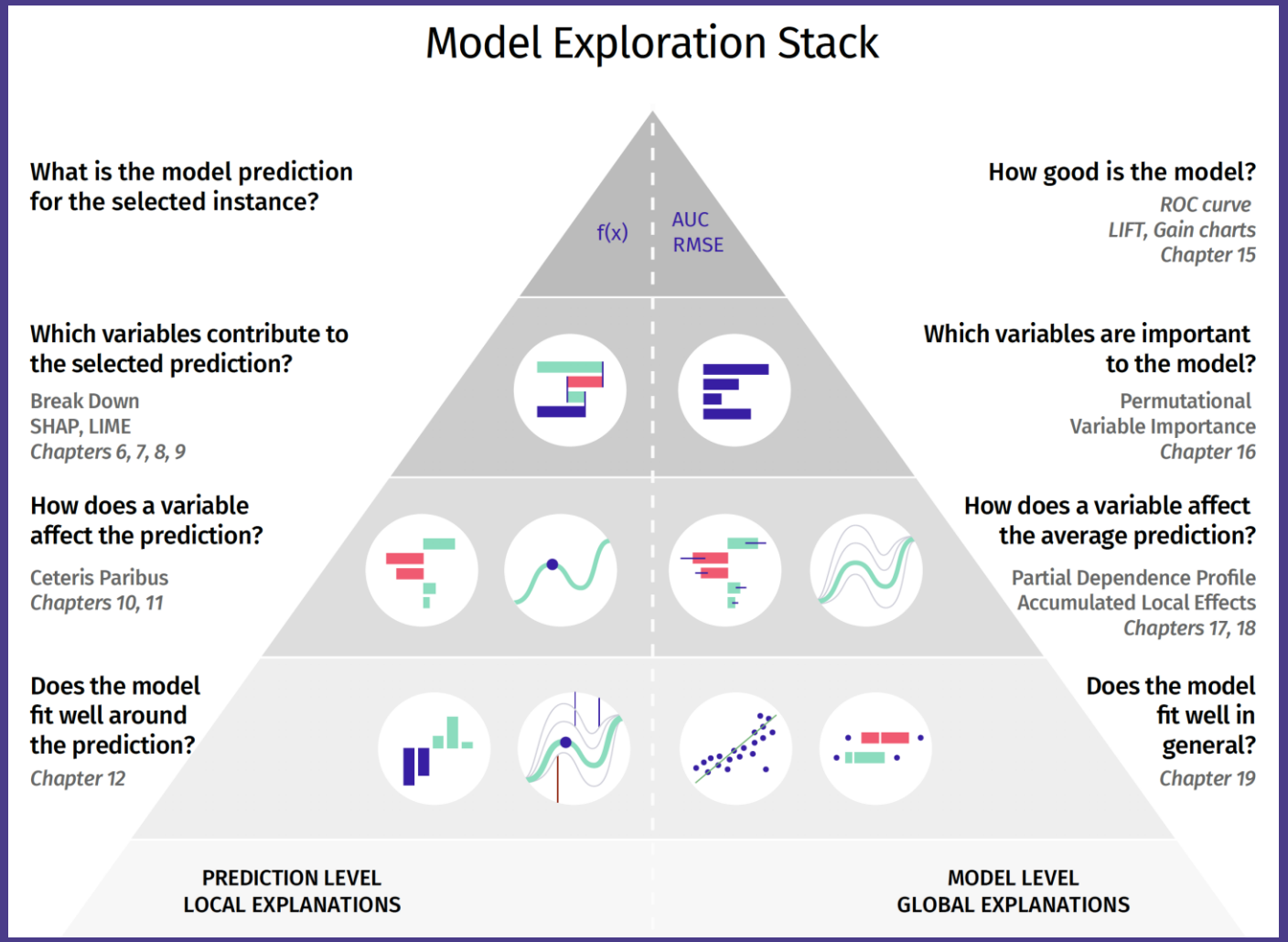
Aim: find a useful tool, resource, idea

What's new?

- **book**: Explanatory Model Analysis
- **tools**: arenar, fairmodels, xai2cloud
- **explanations**: vivo, triplot, treeshap
- **blog**: Responsible ML
- dalex in **Python** 

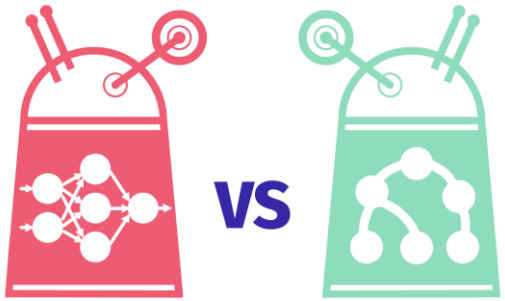


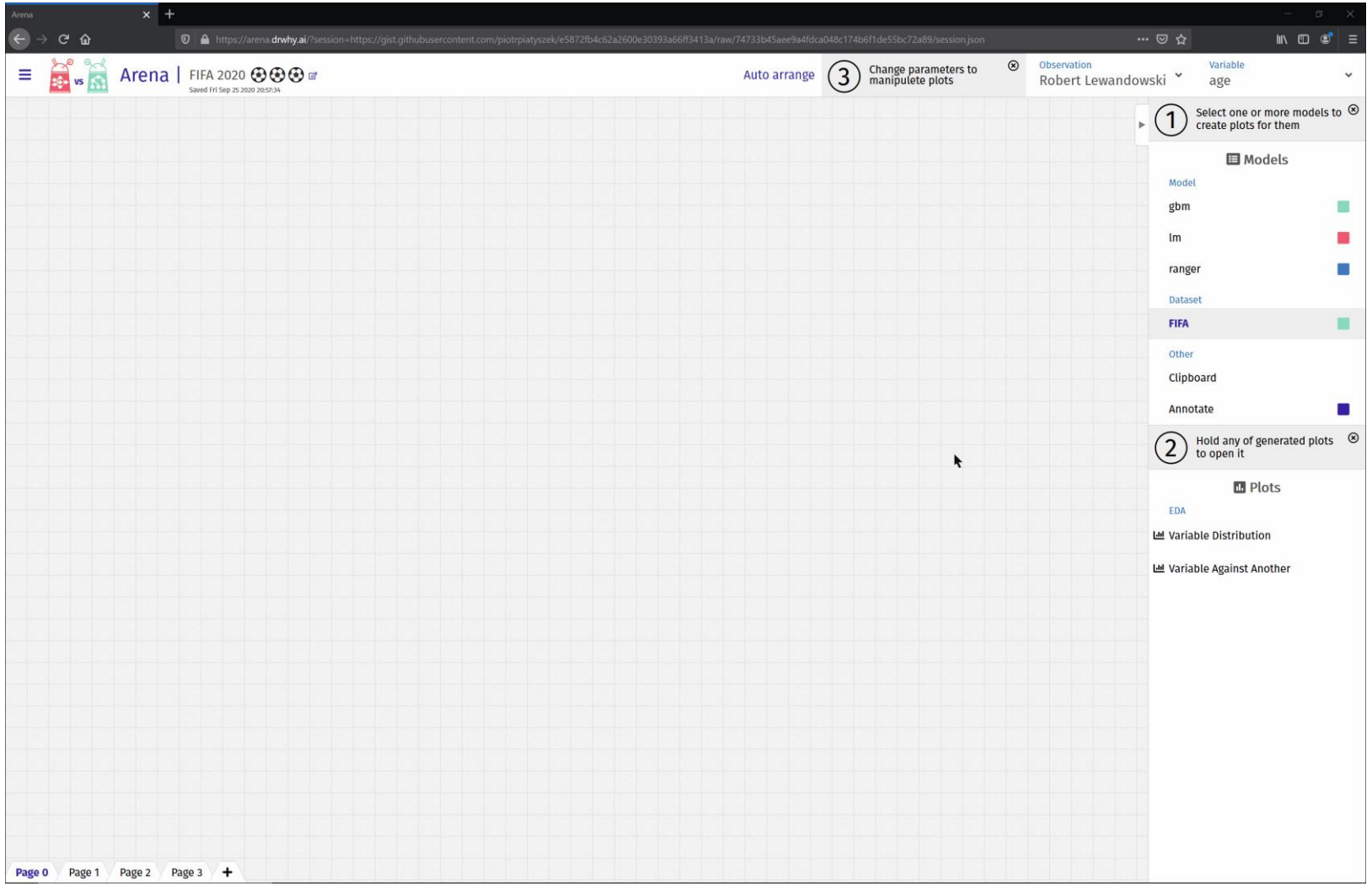
https://pbiecek.github.io/ema/



arenar

- creates a dashboard for Explanatory Model Analysis
- interactive model and data explanation and comparison
- simple to use package
- save & share your analysis

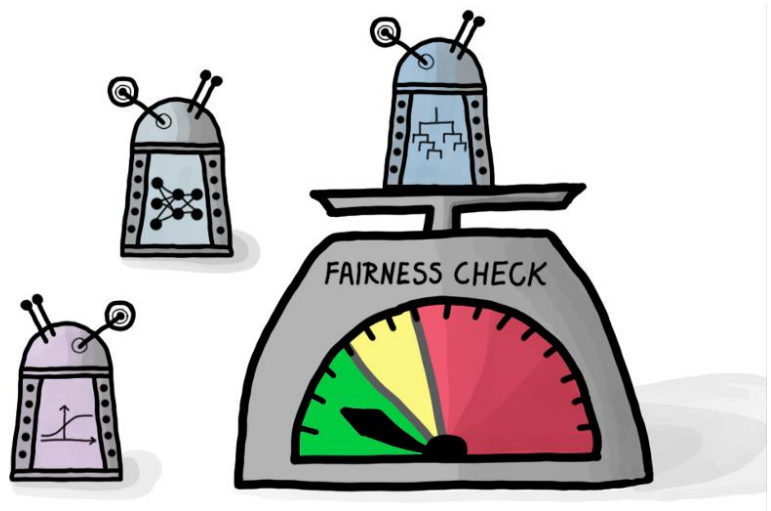


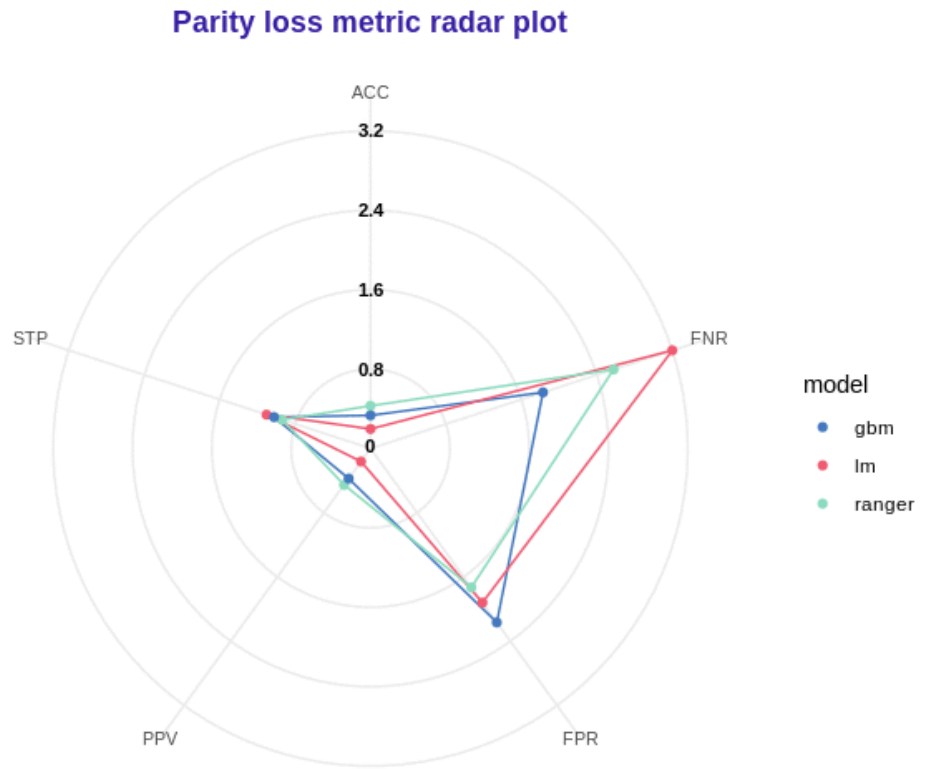


https://github.com/ModelOriented/arenar

fairmodels

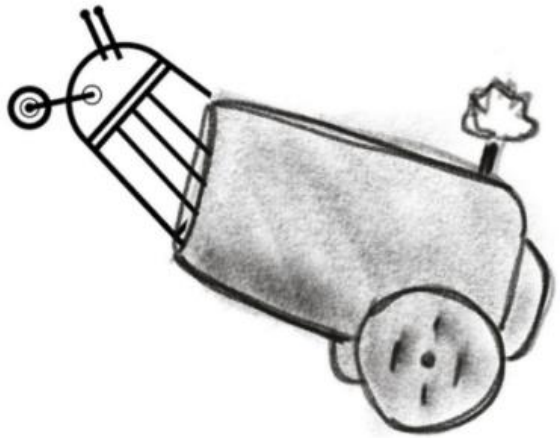
- check model fairness in respect to sensitive categorical variables
- pre- and post- bias mitigation
- compare measures for multiple models
- various techniques and visualisations





https://github.com/ModelOriented/fairmodels

xai2cloud



- automate the deployment of DALEX explainer into the cloud
- compute explanations and retrieve the results with REST API

POST `/predict`

Details

Response body


```
{
  "result_text": [
    "Predicted value 0.4339425057145l8"
  ],
  "result": [
    0.4339
  ],
  "raw_body": []
}
```

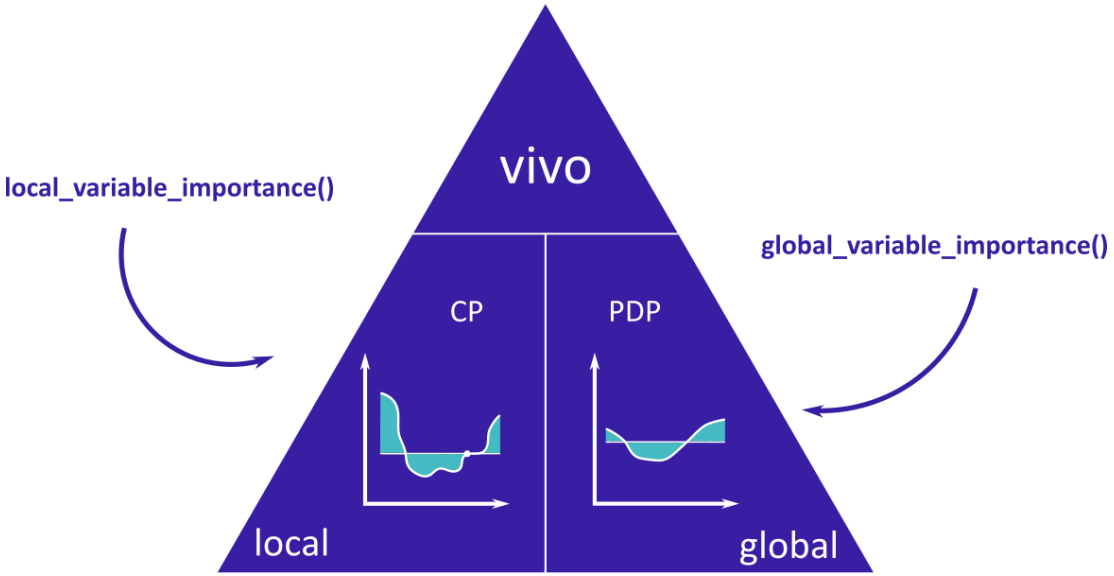

https://github.com/ModelOriented/xai2cloud

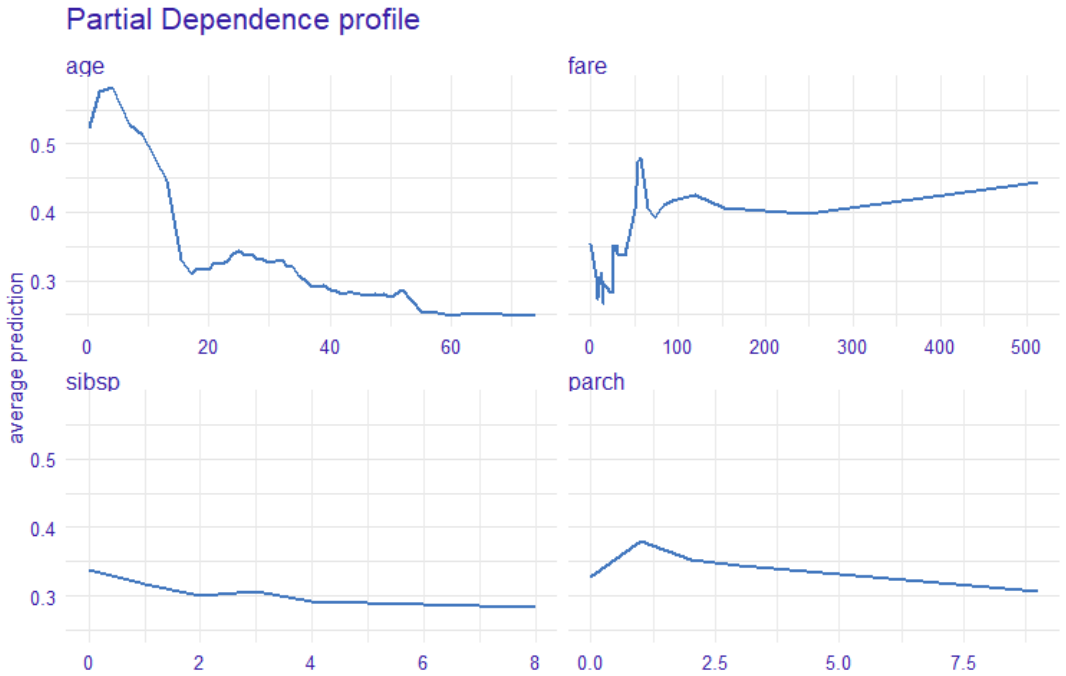
MI

Explanations

vivo

- alternative, model-agnostic way of calculating variable importance based on the Ceteris Paribus and Partial Dependence Profiles
- no random component





https://github.com/ModelOriented/vivo

Explanations

triplot

- a lot of variables ⟶ a lot of correlated variables
- importance of groups of correlated variables – global and local

https://github.com/ModelOriented/triplot

Explanations

treeshap (beta)

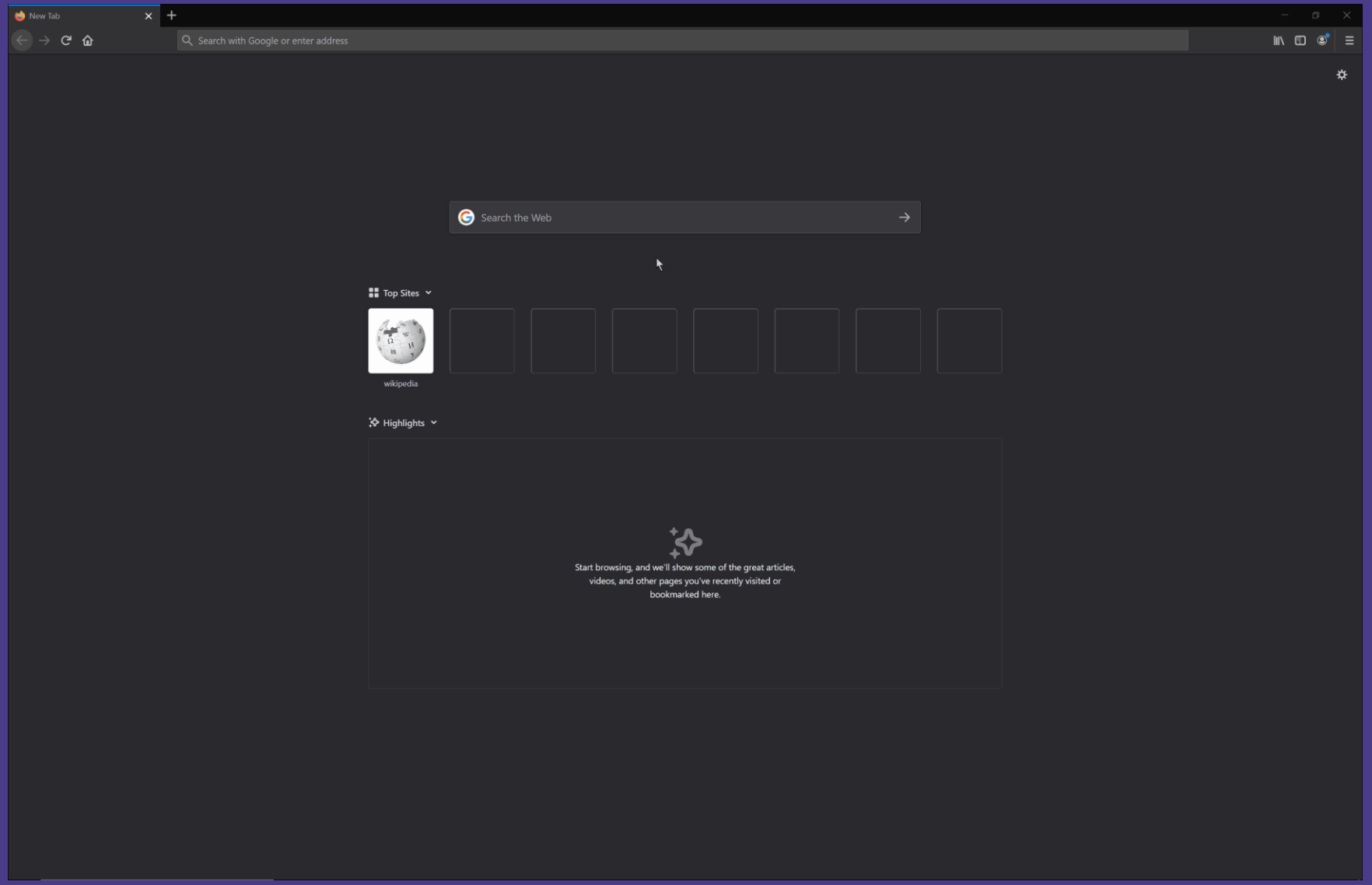
- fast C++ implementation of SHAP values computation in R for tree ensemble models
- global and local plots
- supports models produced with:
  - xgboost, lightgbm, catboost, gbm
  - ranger, randomForest
  - propose more!

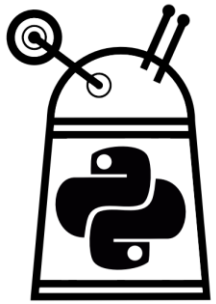
https://github.com/ModelOriented/treeshap

MI

DrWhy.AI blog: Responsible ML

https://medium.com/responsibleml





DALEX

moDel Agnostic Language for Exploration and eXplanation

Python ☆ 649 97

Python-check passing | pypi package 0.2.2 | downloads 11k

pip install dalex

import dalex as dx

MODEL

- scikit-learn
- tensorflow, keras
- xgboost, lightgbm
- ANY

DATA

- pandas
- numpy

dx.Explainer

METHODS

predict/model + parts/profile/diagnostics/surrogate/performance

EXPLANATIONS

- result attribute (pandas)
- plot method (plotly)

Share & send feedback

GitHub	https://github.com/ModelOriented
ResponsibleML	https://medium.com/responsibleml
Contact me	https://linkedin.com/in/hbaniecki