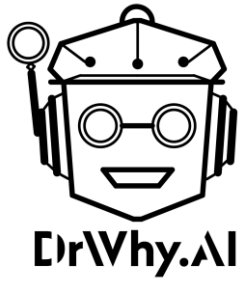
XAI to support prediction making in COVID-19 pandemic

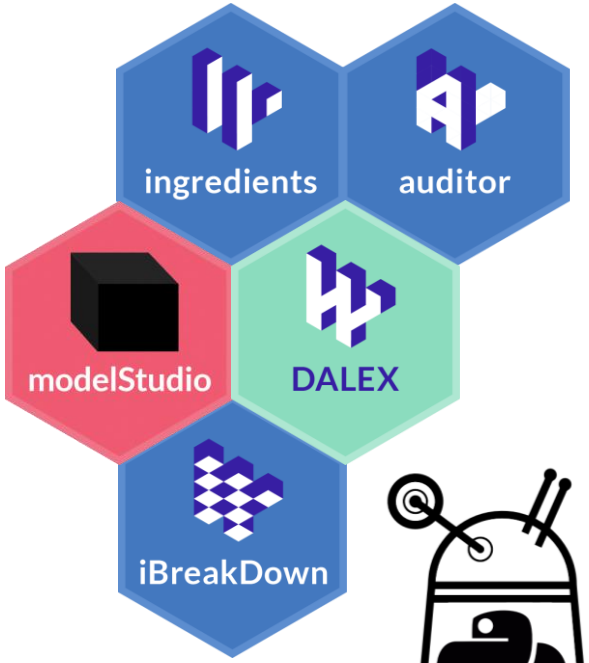
Hubert Baniecki, Warsaw, 10.2020

Hi!

- Data Science student at Warsaw University of Technology
- **Research Software Engineer**

  at MI2 DataLab lead by Przemyslaw Biecek
- interested in XAI and model-human interaction
- creating & maintaining the DrWhy.AI universe





- **member of DeCoviD project** (grant: IDUB against COVID-19)
- **MI2 DataLab in MOCOS group** (MOdelling COronavirus Spread)

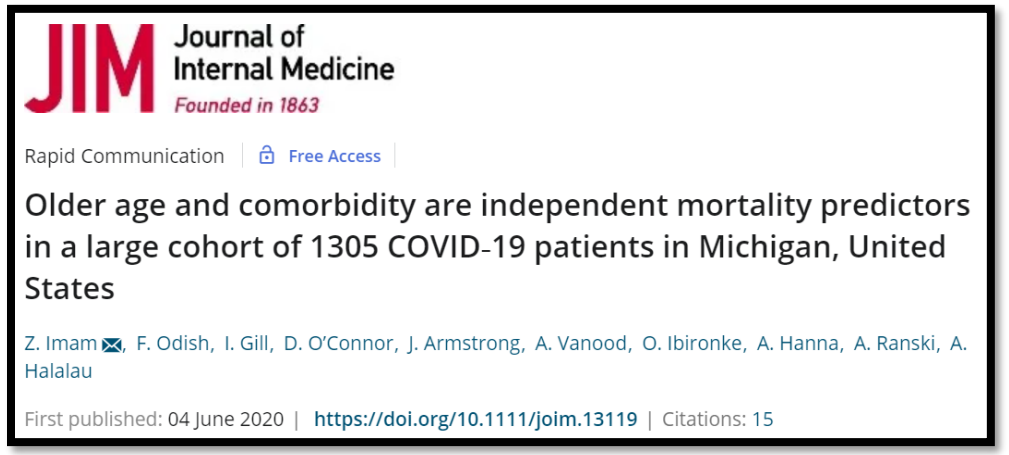
1. Prediction Making in COVID-19
   literature review & critique
2. Real case study
   R & Python XAI tools
3. MI2 DataLab in MOCOS
   COVID-19 analysis for Poland

1.Prediction Making in COVID-19

literature review & critique

Literature

- diagnosis of COVID-19
- prognosis of patient mortality risk
- prediction of ilness severity, hospitalization risk, length of hospital stay
- tabular data (e.g. age, blood test) & image data (medical lung images)
- statistical models / ML models / DL models



JIM Journal of Internal Medicine
Founded in 1863

Rapid Communication | Free Access

Older age and comorbidity are independent mortality predictors in a large cohort of 1305 COVID-19 patients in Michigan, United States

Z. Imam, F. Odish, I. Gill, D. O'Connor, J. Armstrong, A. Vanood, O. Ibironke, A. Hanna, A. Ranski, A. Halalau

First published: 04 June 2020 | https://doi.org/10.1111/joim.13119 | Citations: 15

Image & Signal Processing | Published: 01 July 2020

Detection of COVID-19 Infection from Routine Blood Exams with Machine Learning: A Feasibility Study

Davide Brinati, Andrea Campagner, Davide Ferrari, Massimo Locatelli, Giuseppe Banfi & Federico Cabitza

Journal of Medical Systems **44**, Article number: 135 (2020) | Cite this article

2609 Accesses | **7** Altmetric | Metrics

Predicting Mortality Due to SARS-CoV-2: A Mechanistic Score Relating Obesity and Diabetes to COVID-19 Outcomes in Mexico FREE

Omar Yaxmehen Bello-Chavolla, Jessica Paola Bahena-López, Neftali Eduardo Antonio-Villa, Arsenio Vargas-Vázquez, Armando González-Díaz, Alejandro Márquez-Salinas, Carlos A Fermín-Martínez, J Jesús Naveja, Carlos A Aguilar-Salinas

The Journal of Clinical Endocrinology & Metabolism, Volume 105, Issue 8, August 2020, Pages 2752–2761, https://doi.org/10.1210/clinem/dgaa346

Published: 31 May 2020 **Article history**

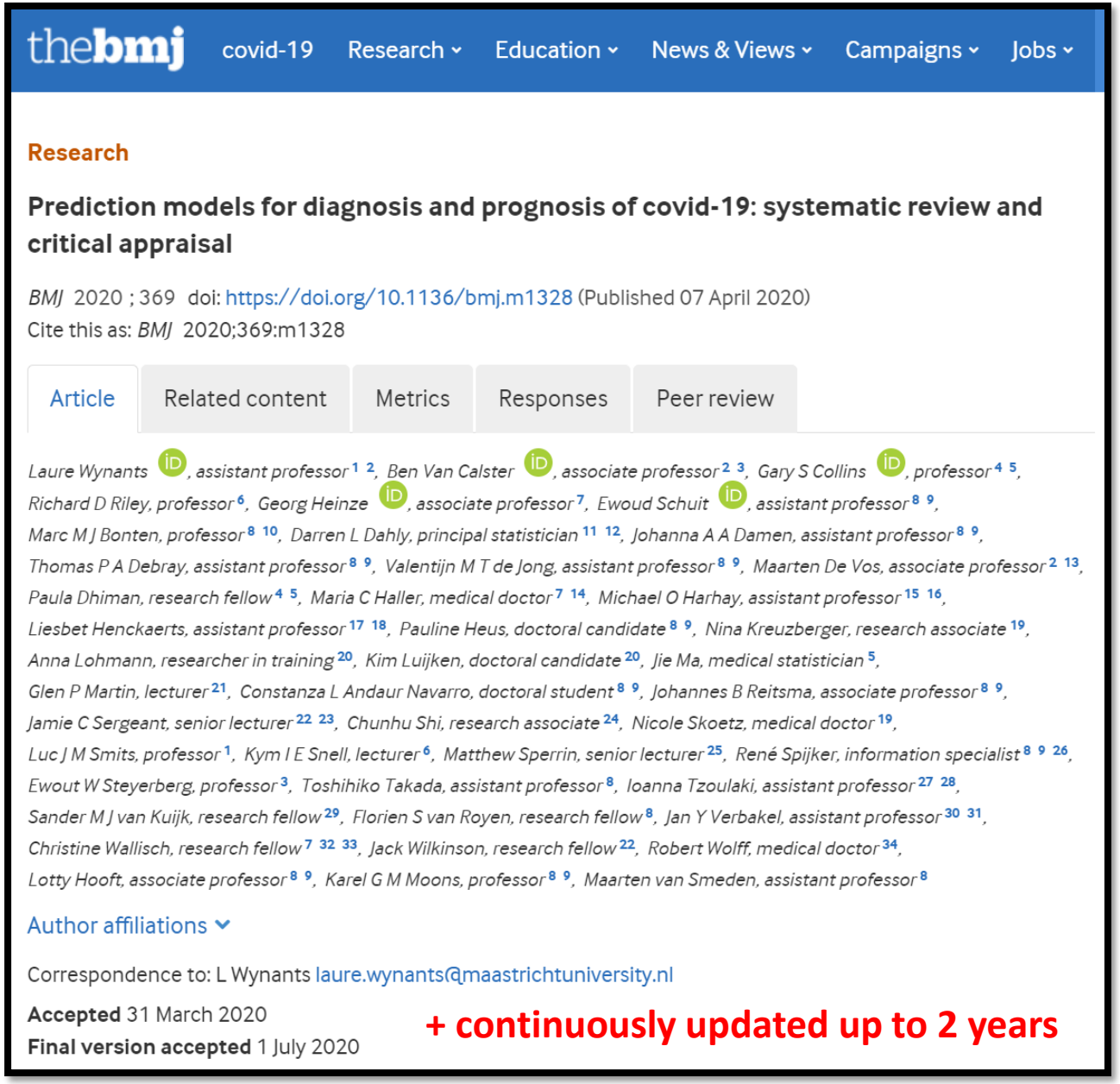
Review

107 studies

91 diagnostic models
2/3 on medical images

50 prognostic models
1/2 for mortality
1/6 for ilness severity

mostly preprints



thebmj covid-19 Research Education News & Views Campaigns Jobs

Research

Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal

BMJ 2020 ; 369 doi: https://doi.org/10.1136/bmj.m1328 (Published 07 April 2020)
Cite this as: *BMJ* 2020;369:m1328

Article | Related content | Metrics | Responses | Peer review

Laure Wynants, assistant professor[1 2], Ben Van Calster, associate professor[2 3], Gary S Collins, professor[4 5], Richard D Riley, professor[6], Georg Heinze, associate professor[7], Ewoud Schuit, assistant professor[8 9], Marc M J Bonten, professor[8 10], Darren L Dahly, principal statistician[11 12], Johanna A A Damen, assistant professor[8 9], Thomas P A Debray, assistant professor[8 9], Valentijn M T de Jong, assistant professor[8 9], Maarten De Vos, associate professor[2 13], Paula Dhiman, research fellow[4 5], Maria C Haller, medical doctor[7 14], Michael O Harhay, assistant professor[15 16], Liesbet Henckaerts, assistant professor[17 18], Pauline Heus, doctoral candidate[8 9], Nina Kreuzberger, research associate[19], Anna Lohmann, researcher in training[20], Kim Luijken, doctoral candidate[20], Jie Ma, medical statistician[5], Glen P Martin, lecturer[21], Constanza L Andaur Navarro, doctoral student[8 9], Johannes B Reitsma, associate professor[8 9], Jamie C Sergeant, senior lecturer[22 23], Chunhu Shi, research associate[24], Nicole Skoetz, medical doctor[19], Luc J M Smits, professor[1], Kym I E Snell, lecturer[6], Matthew Sperrin, senior lecturer[25], René Spijker, information specialist[8 9 26], Ewout W Steyerberg, professor[3], Toshihiko Takada, assistant professor[8], Ioanna Tzoulaki, assistant professor[27 28], Sander M J van Kuijk, research fellow[29], Florien S van Royen, research fellow[8], Jan Y Verbakel, assistant professor[30 31], Christine Wallisch, research fellow[7 32 33], Jack Wilkinson, research fellow[22], Robert Wolff, medical doctor[34], Lotty Hooft, associate professor[8 9], Karel G M Moons, professor[8 9], Maarten van Smeden, assistant professor[8]

Author affiliations

Correspondence to: L Wynants laure.wynants@maastrichtuniversity.nl

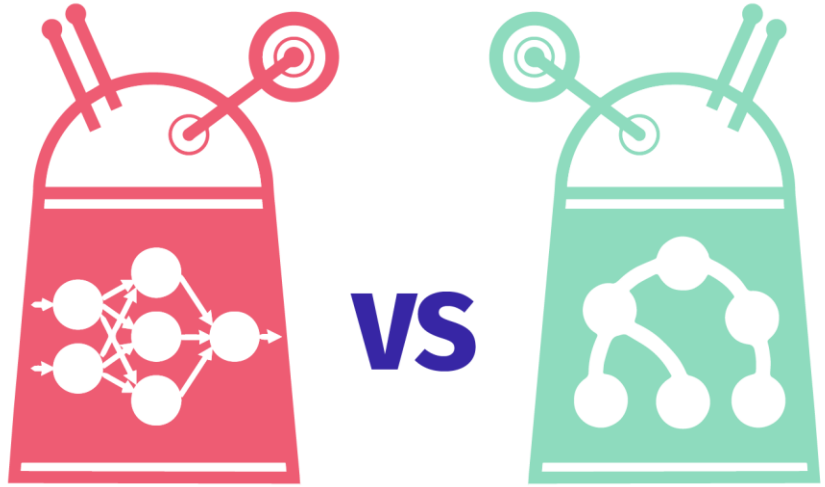
Accepted 31 March 2020
Final version accepted 1 July 2020

+ continuously updated up to 2 years

https://www.bmj.com/content/369/bmj.m1328

Critique

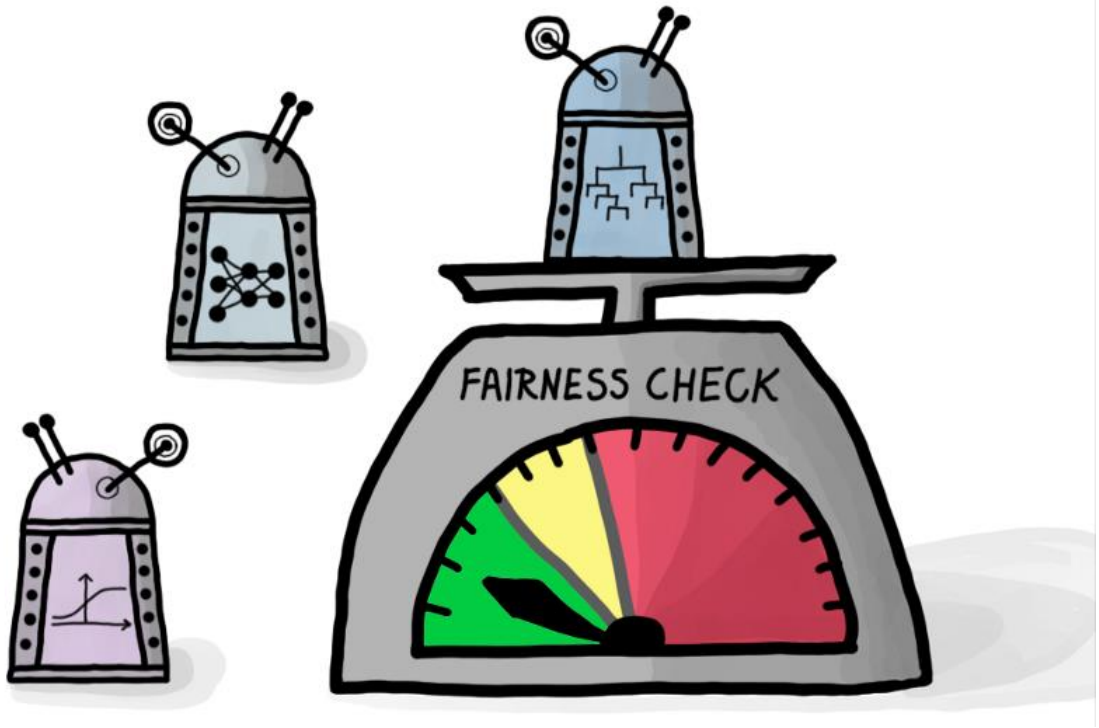
- all models biased and/or overfitted
- insufficiently validated
- only 1/3 with code!
- limited, poorly diversified data
- do not use these models in practice
- future: better validation for prediction making



https://www.bmj.com/content/369/bmj.m1328

Responsible AI, ML, Be Responsible!

- Explainable AI, nowadays Responsible AI
- data exploration & model explanation
- fair, unbiased models
- complete model documentation
- performance is not validation



2.Real case study

R & Python XAI tools

Case study

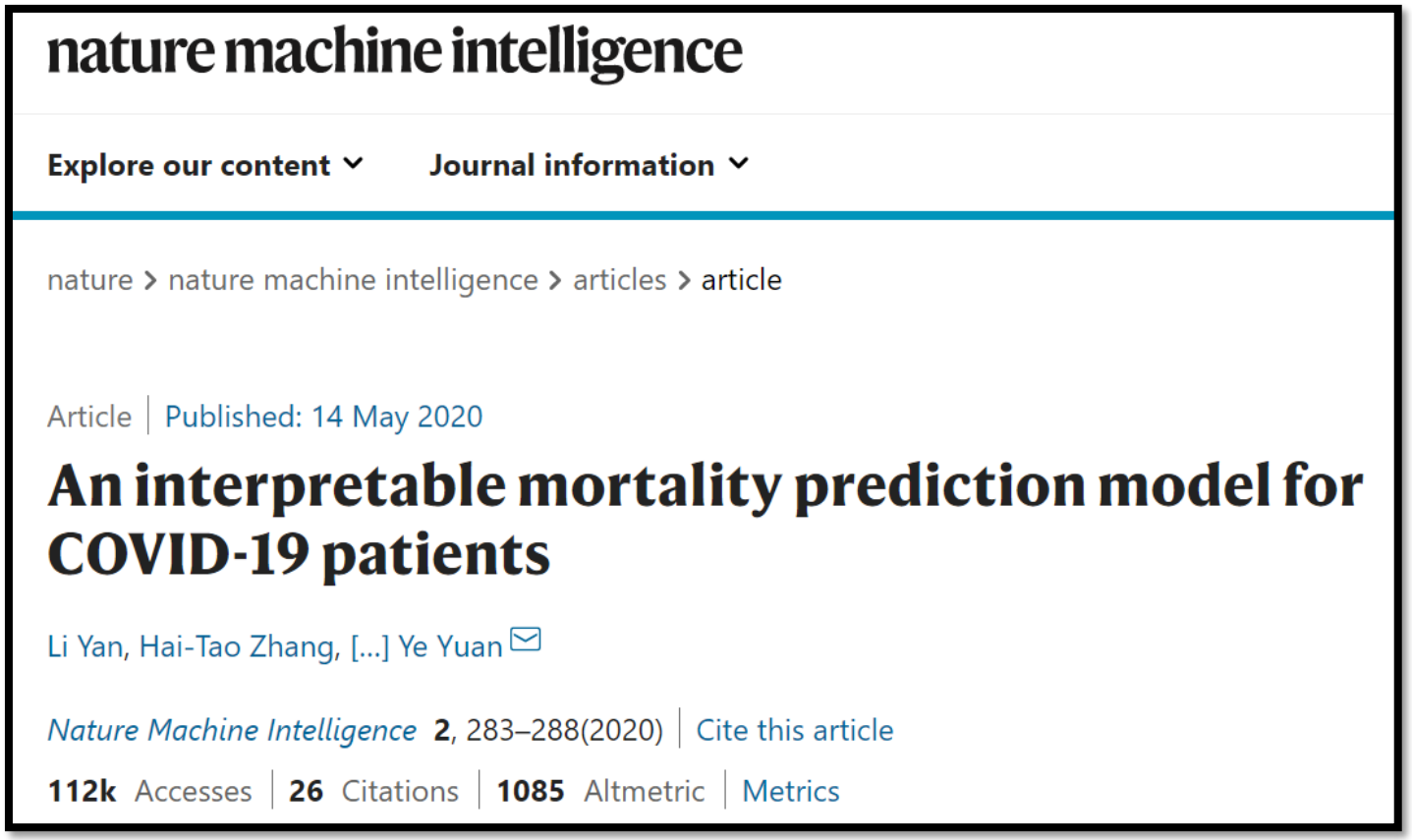
February, Wuhan, China

COVID-19 infected patients
485 (375 + 110)
external test dataset

predict mortality from clinical variables (e.g. blood test)

XGBoost (0.99 AUC) to assess variable importance

interpretable decision tree
3 variables – 0.96 AUC



nature machine intelligence

Explore our content ⌄ Journal information ⌄

nature › nature machine intelligence › articles › article

Article | Published: 14 May 2020

An interpretable mortality prediction model for COVID-19 patients

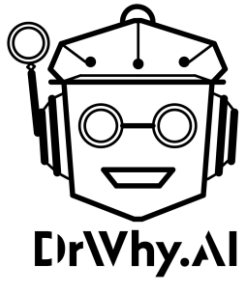
Li Yan, Hai-Tao Zhang, [...] Ye Yuan ✉

Nature Machine Intelligence **2**, 283–288(2020) | Cite this article

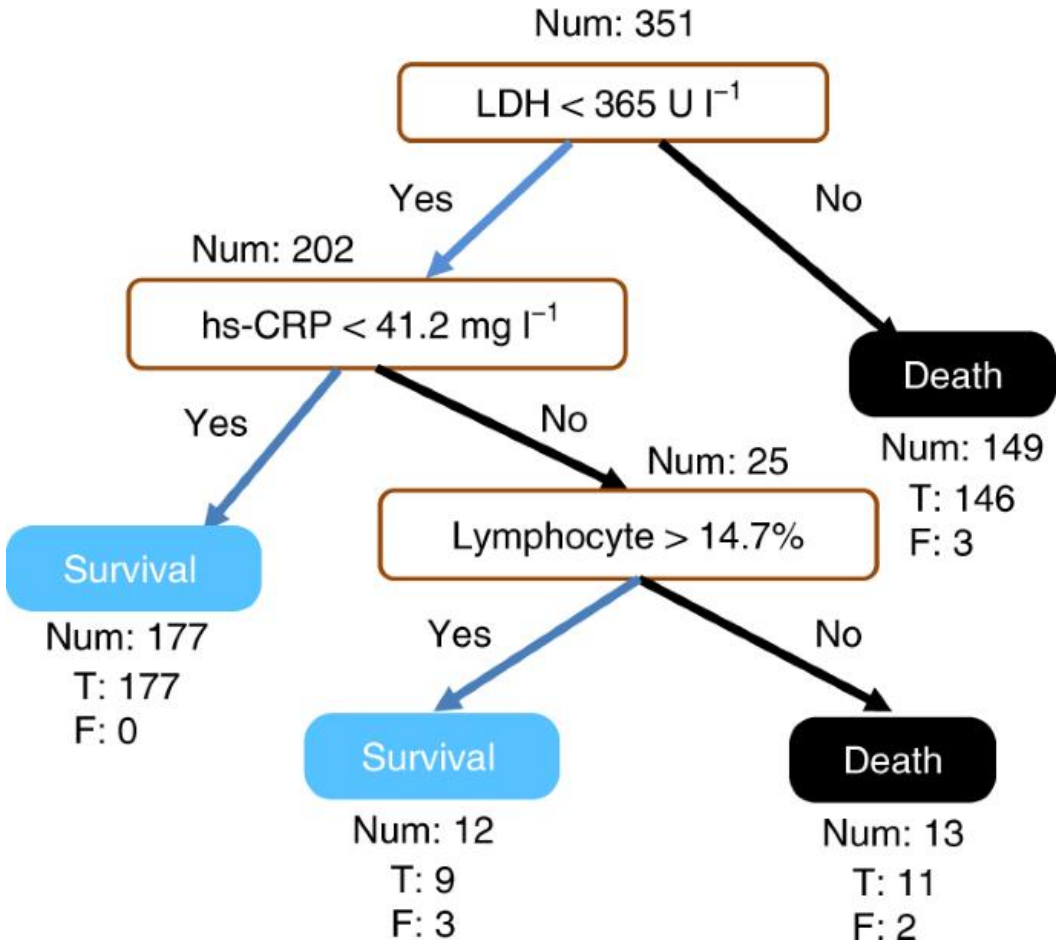
112k Accesses | **26** Citations | **1085** Altmetric | Metrics

https://www.nature.com/articles/s42256-020-0180-7

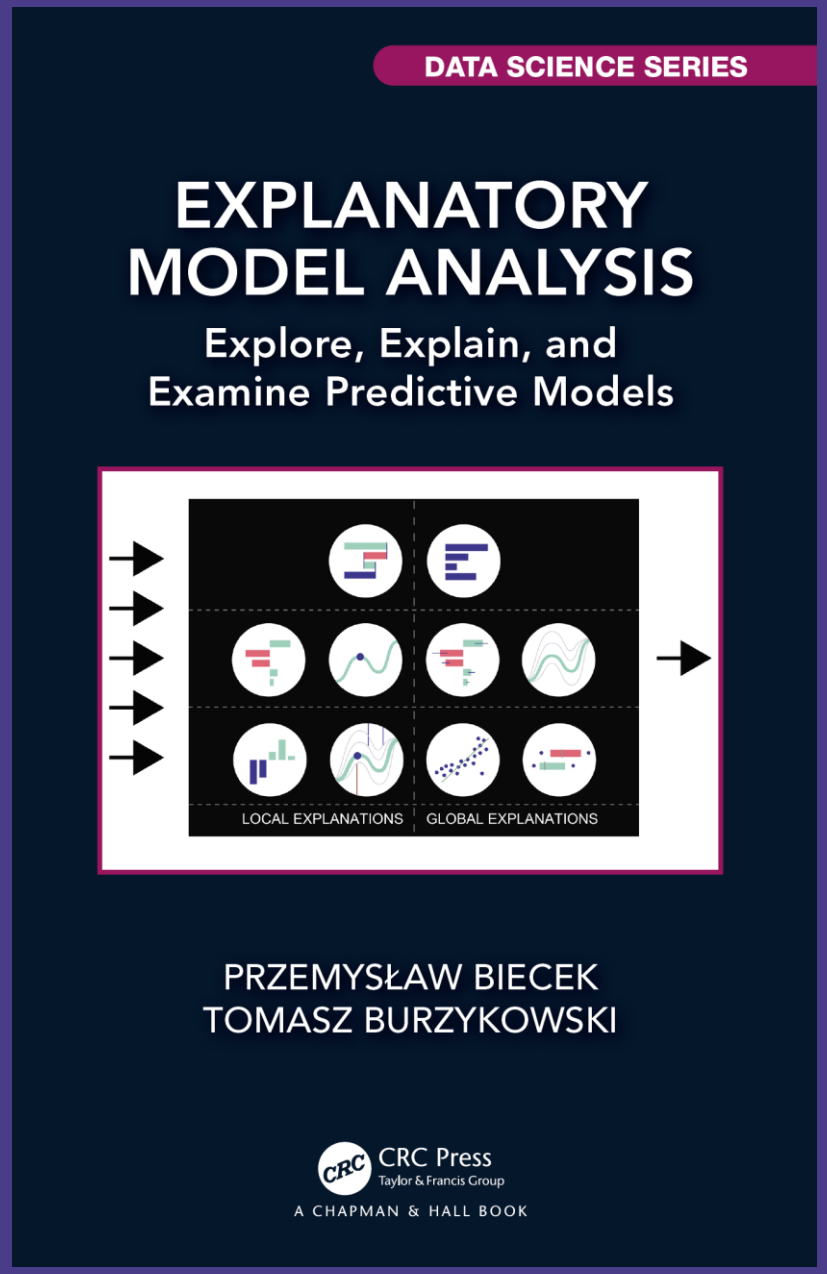
Wh-questions



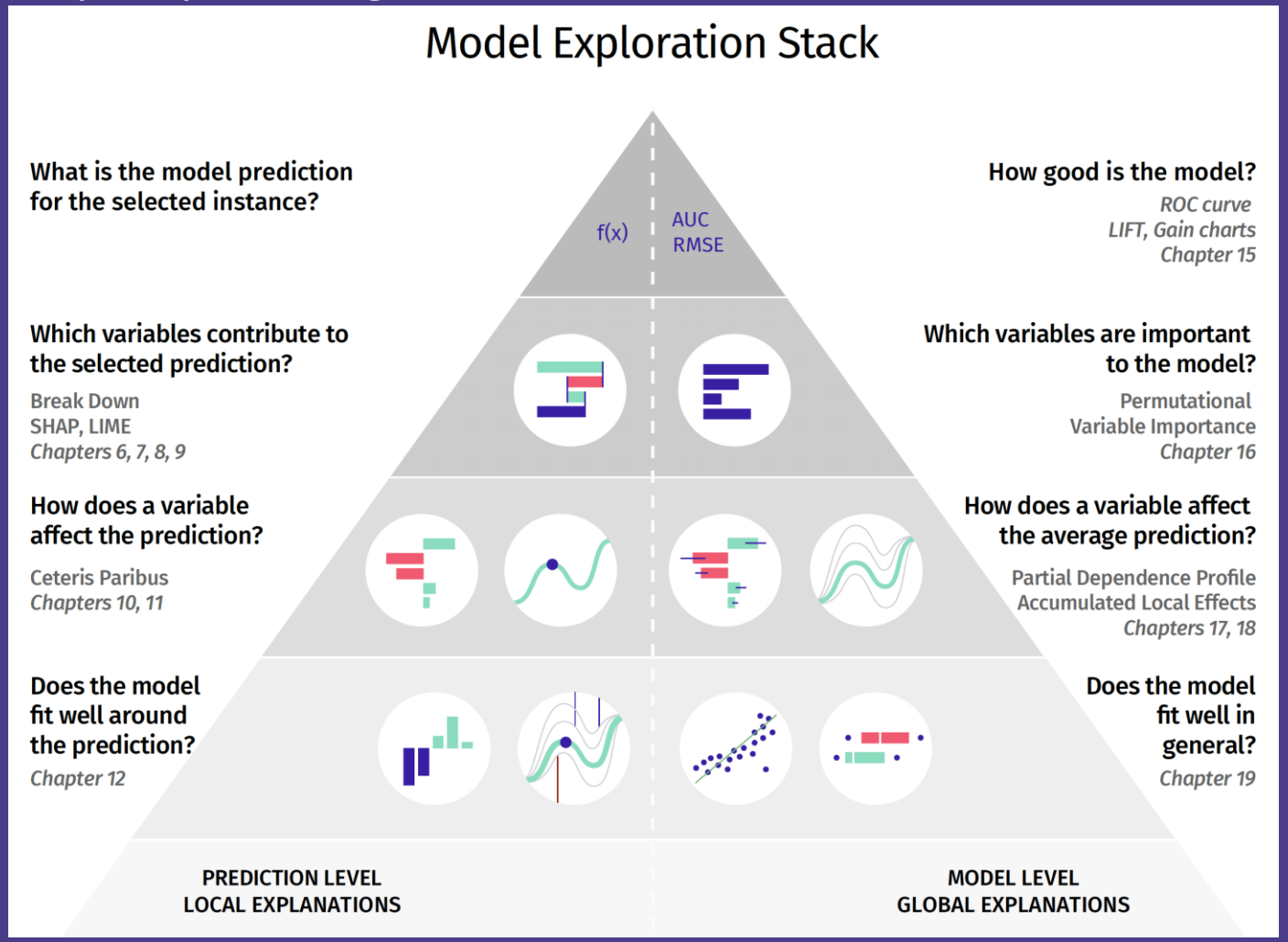
- Why is the performance so good?
- Why 3 variables out of 50 are sufficient?
- Why was age not important?
- What are the continuous relationships between variables and the target in the XGBoost model?
- …

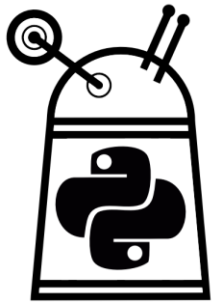


https://www.nature.com/articles/s42256-020-0180-7



https://pbiecek.github.io/ema





DALEX

moDel Agnostic Language for Exploration and eXplanation

Python ☆ 649 97

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

MODEL

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

DATA

- pandas
- numpy

pip install dalex

import dalex as dx

dx.Explainer

METHODS

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

EXPLANATIONS

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

jupyter DSS2020 Last Checkpoint: a minute ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted Python 3

Run Code

In [ ]:
```
import seaborn as sns
import xgboost as xgb
import dalex as dx
```

In [ ]:
```
# paper supplementary code
from utils_features_selection import data_preprocess, col_miss
```

In [ ]:
```
def data_read():
    """
    read paper supplementary data and preprocess
    """

    data_df_unna, data_pre_df = data_preprocess()
    col_miss_data = col_miss(data_df_unna)
    col_miss_data['Missing_part'] = col_miss_data['missing_count'] / len(data_df_unna)
    sel_cols = col_miss_data[col_miss_data['Missing_part'] <= 0.2]['col']
    data_df_sel = data_df_unna[sel_cols].copy()
    data_raw = data_df_sel.fillna(-1)
    name_dict = {'乳酸脱氢酶': 'LDH', '淋巴细胞(%)': 'Lymphocytes(%)', '超敏C反应蛋白': 'hs-CRP',
                 '年龄': 'Age', 'Type2': 'Target'}
    data_raw.rename(columns=name_dict, inplace=True)
    data_train = data_raw.loc[:, name_dict.values()]
    return data_train
```

In [ ]:
```
# load data
data = data_read()

# change Target [0.0, 1.0] to [0, 1]
data.Target = data.Target.astype(int)

# plot data exploration
colors = ["#4378bf", "#f05a71"]
sns.set_palette(sns.color_palette(colors))
g = sns.pairplot(data, hue='Target', corner=True)
g._legend.set_bbox_to_anchor((0.75, 0.75))
g.fig.set_size_inches(10.5, 10)
```

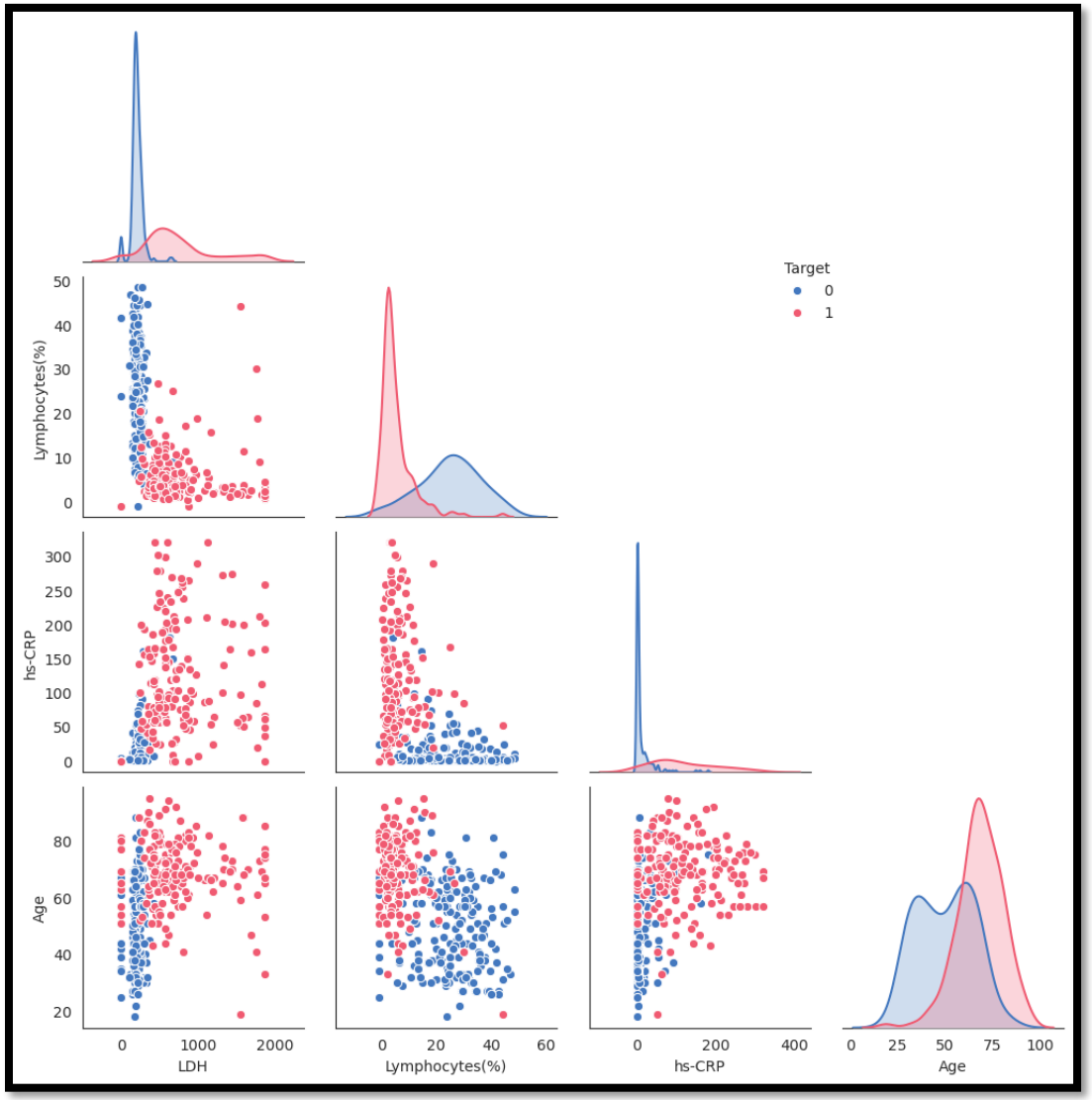
In [ ]:
```
# load data
data = data_read()
X, y = data.drop('Target', axis=1), data.Target

# make a model
model = xgb.XGBClassifier(max_depth=4, learning_rate=0.2, reg_lambda=1,
                          n_estimators=150, subsample=0.9, colsample_bytree=0.9,
                          random_state=0)
model.fit(X, y)
```


https://github.com/hbaniecki/Pre-Surv-COVID-19

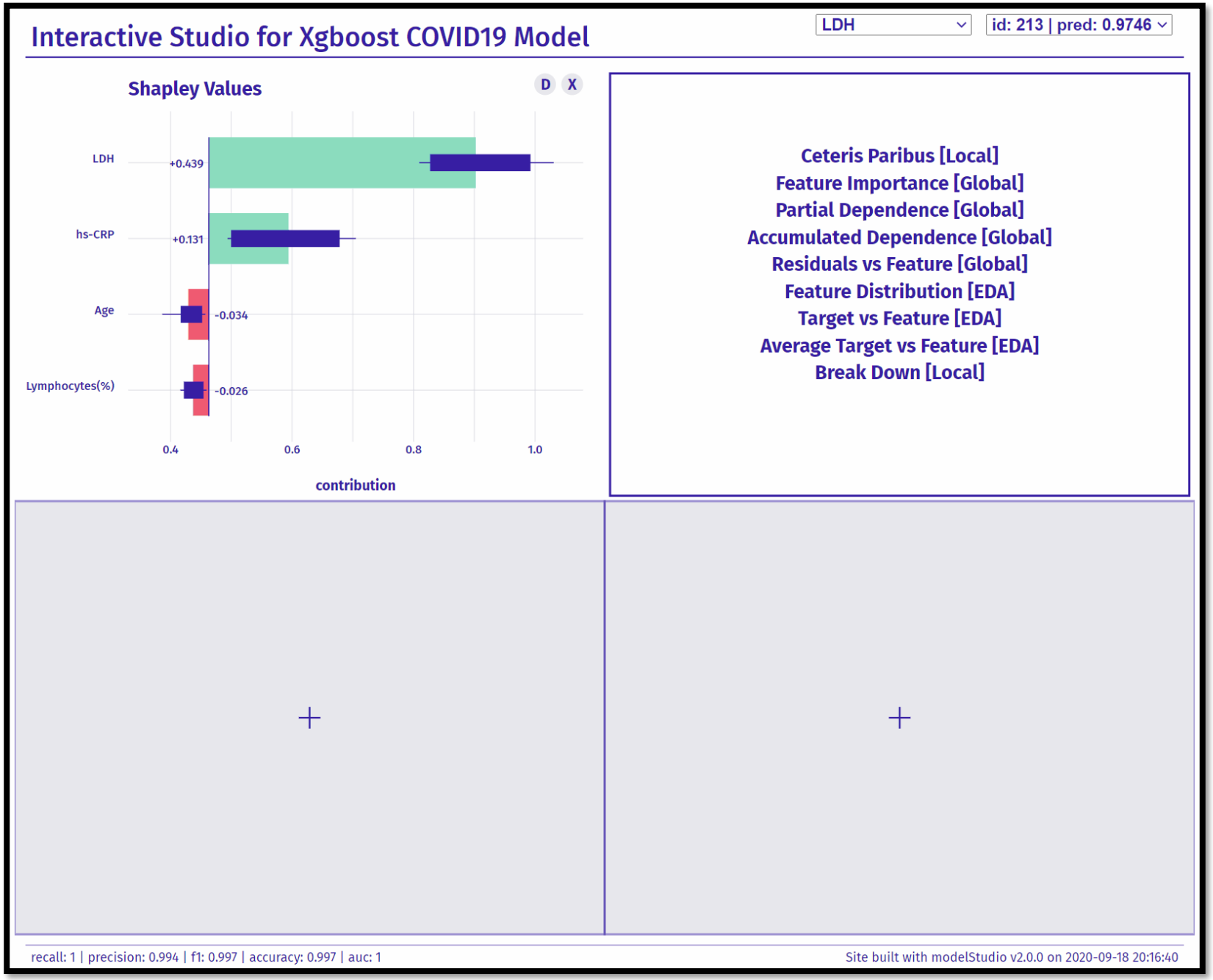
MI

Data exploration



	Accuracy	Balanced Accuracy	ROC AUC	F1 Score	Time Taken
Model					
LGBMClassifier	1.00	1.00	1.00	1.00	0.03
XGBClassifier	1.00	1.00	1.00	1.00	0.03
RandomForestClassifier	1.00	1.00	1.00	1.00	0.16
ExtraTreesClassifier	0.99	0.99	0.99	0.99	0.15
SVC	0.99	0.99	0.99	0.99	0.01
KNeighborsClassifier	0.99	0.99	0.99	0.99	0.03
GaussianNB	0.99	0.99	0.99	0.99	0.02
AdaBoostClassifier	0.99	0.99	0.99	0.99	0.09
QuadraticDiscriminantAnalysis	0.98	0.98	0.98	0.98	0.02
NuSVC	0.97	0.97	0.97	0.97	0.02
LinearSVC	0.97	0.97	0.97	0.97	0.01
LinearDiscriminantAnalysis	0.97	0.97	0.97	0.97	0.02
BaggingClassifier	0.97	0.97	0.97	0.97	0.04
LogisticRegression	0.97	0.97	0.97	0.97	0.02
CalibratedClassifierCV	0.97	0.97	0.97	0.97	0.13
RidgeClassifier	0.97	0.97	0.97	0.97	0.02
RidgeClassifierCV	0.97	0.97	0.97	0.97	0.01
DecisionTreeClassifier	0.96	0.96	0.96	0.96	0.04
Perceptron	0.96	0.96	0.96	0.96	0.02
SGDClassifier	0.96	0.96	0.96	0.96	0.01
NearestCentroid	0.96	0.96	0.96	0.96	0.02
LabelPropagation	0.95	0.94	0.94	0.95	0.02
LabelSpreading	0.95	0.94	0.94	0.95	0.02
BernoulliNB	0.95	0.94	0.94	0.95	0.02

https://github.com/hbaniecki/Pre-Surv-COVID-19

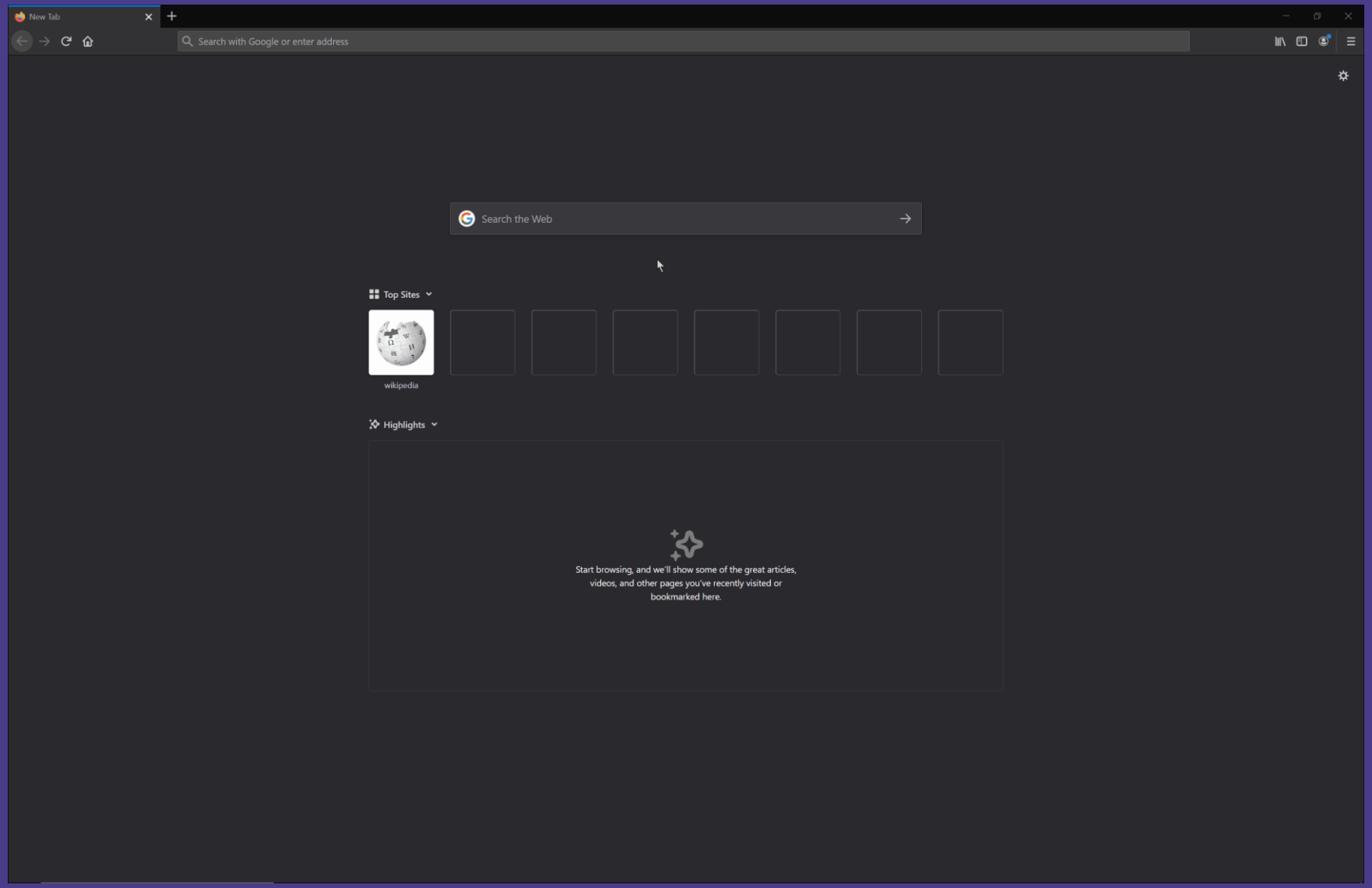


vs **AUC, ACC**
PREDICTION

https://github.com/hbaniecki/Pre-Surv-COVID-19

DrWhy.AI blog: Responsible ML

https://medium.com/responsibleml

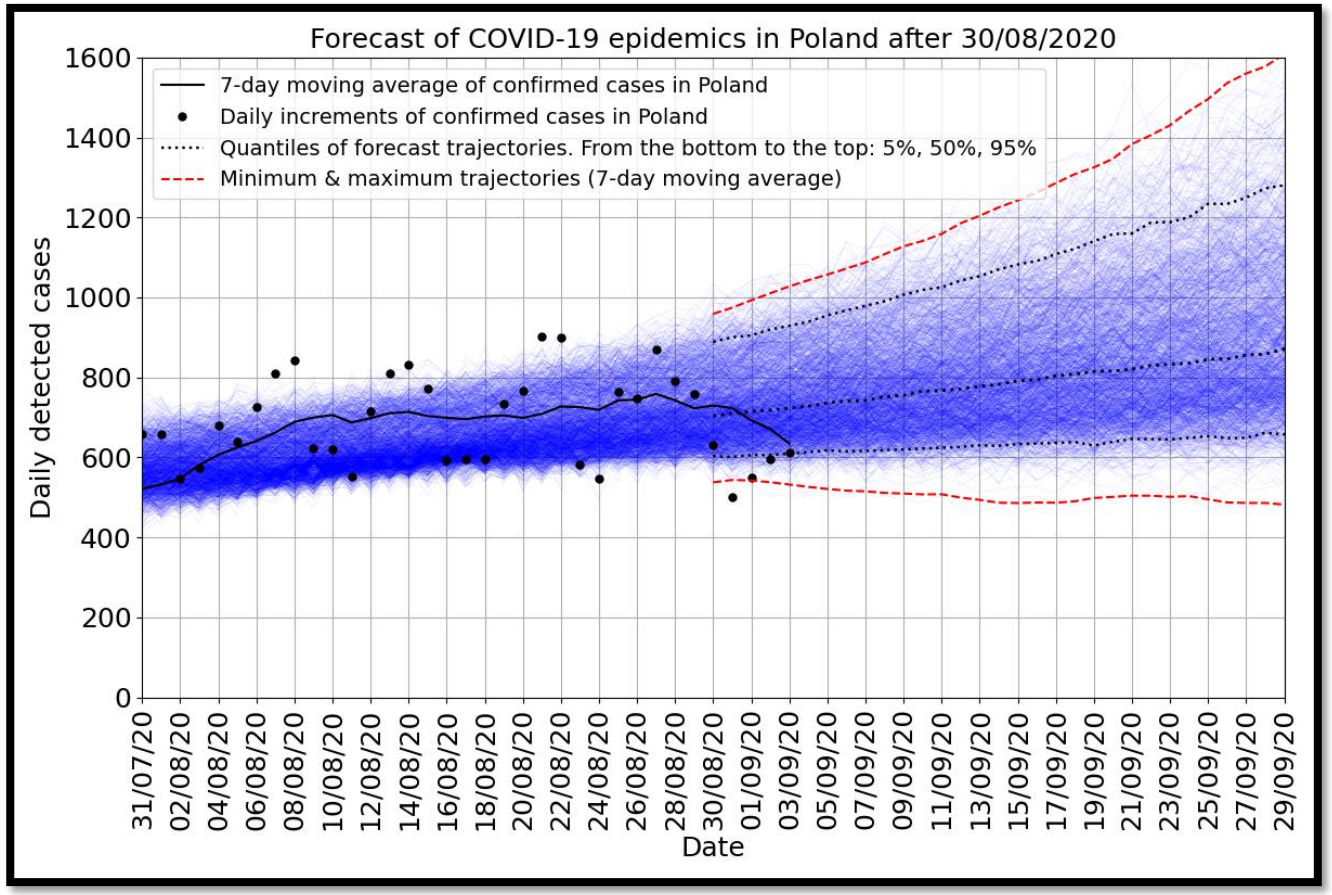


3.MI2 DataLab in MOCOS

COVID-19 analysis for Poland

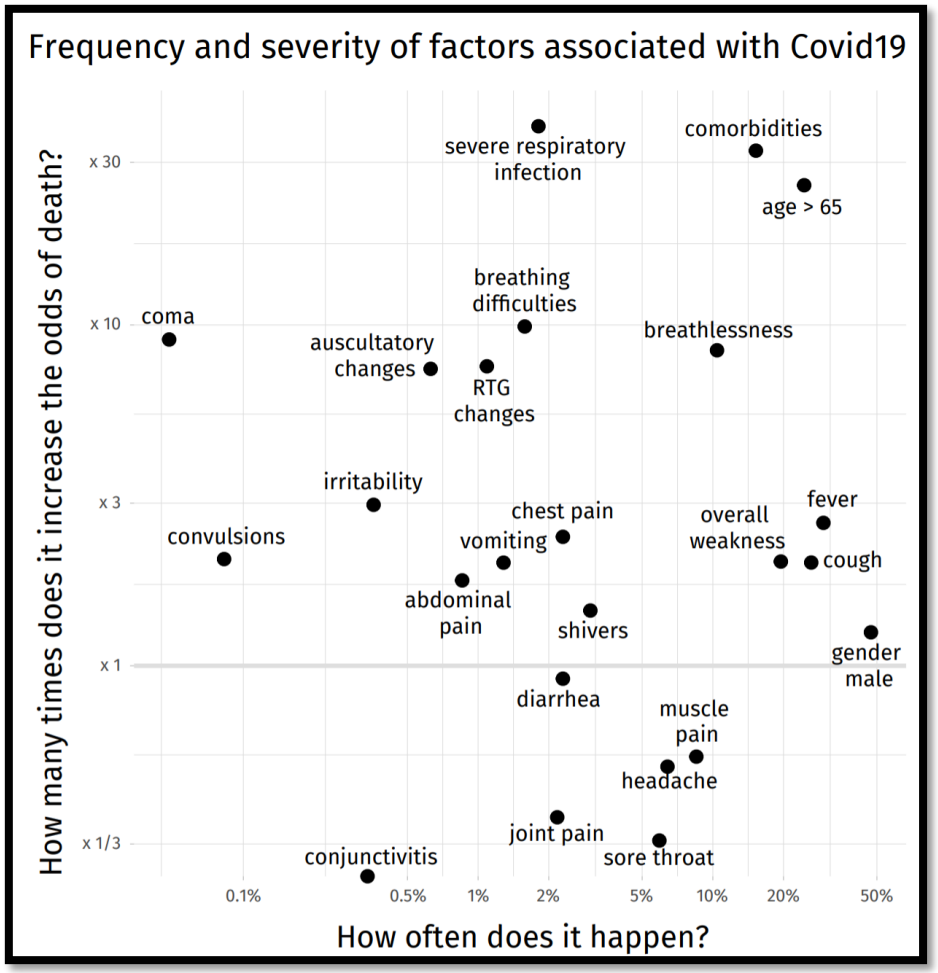
COVID-19 in Poland

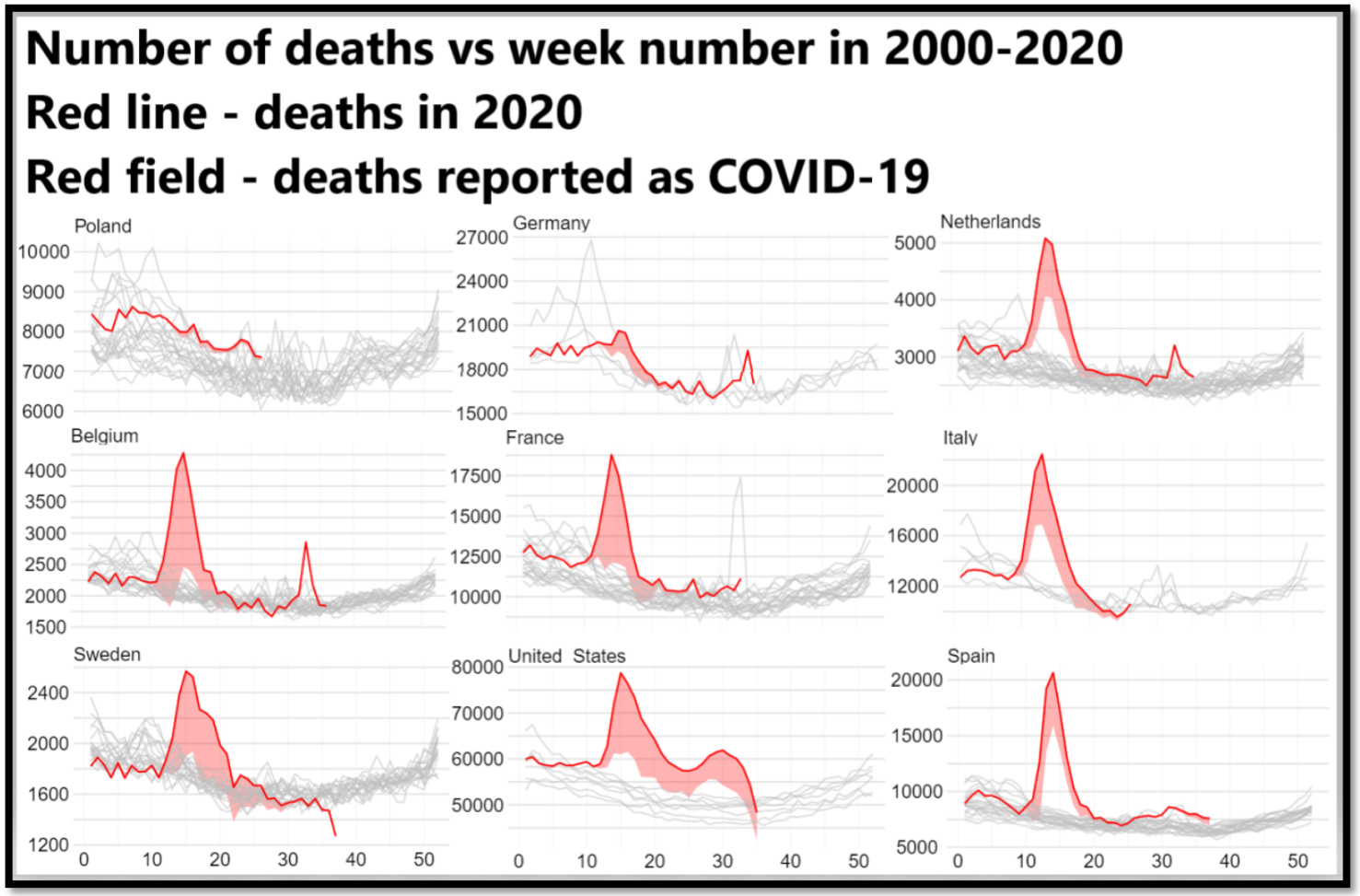
- **MI2 DataLab in MOCOS**
- MOdelling COronavirus Spread group lead by Tyll Krüger
- specific data of 30k COVID-19 infected people in Poland
- reports to Ministry of Health
- data exploration, model development & explanation



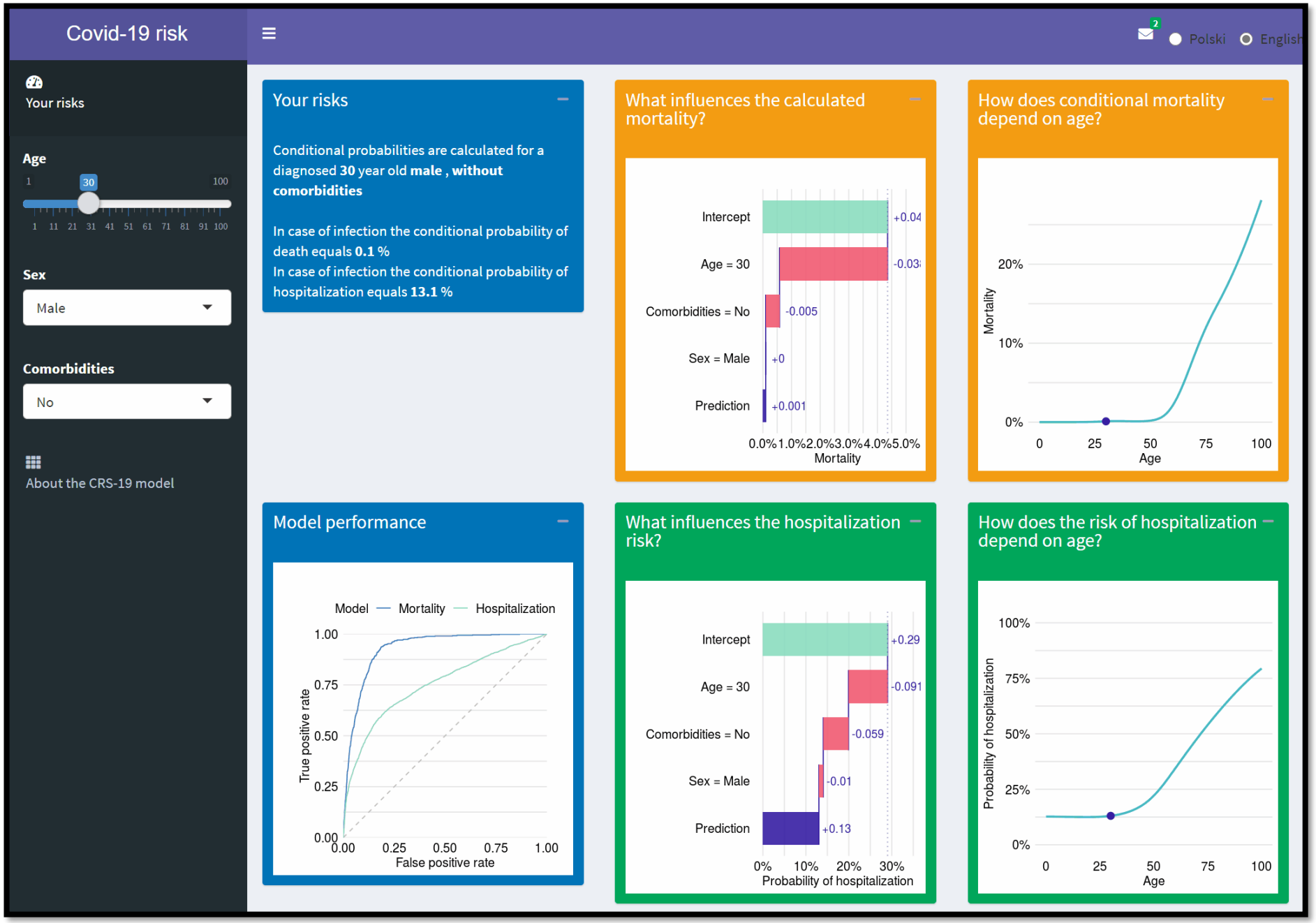
https://mocos.pl/index_en.html

Data exploration





https://mocos.pl/index_en.html



vs

AUC, ACC
PREDICTION

https://mocos.pl/risk_en.html

Takeaways

Advance the prediction making (not only in COVID-19 pandemic):

- Be Responsible!
- Performance-based model validation is not enough
- Proper results need data exploration & model explanation
- **DrWhy.AI:** R & Python XAI tools, resources for Responsible ML

Contact me

@hbaniecki	https://linkedin.com/in/hbaniecki
MI2 DataLab	https://linkedin.com/company/mi2datalab
DrWhy.AI	https://drwhy.ai
ResponsibleML	https://medium.com/responsibleml