Identifying a Trial Population for Clinical Studies on Diabetes Drug Testing with Neural Networks

Tim Löhr
Supervisor: Philipp Schlieper
Machine Learning and Data Analytics for Industry 4.0
Final Presentation
Machine Learning and Data Analytics (MaD) Lab
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)
July 15, 2020

Motivation

Use more
Machine Learning
in Medicine.

To save more lifes.

Figure 1. AI https://www.hhmglobal.com/industry-updates/press-releases/artificial-intelligence-ai-medicine-in-a-block-chain-device

Pave the way

The new

Electronic Patient File

coming into force

on 1.1.2021

Jens Spahn
German Minister
of Health

Figure 2. EHR https://de.cleanpng.com/png-chs524/download-png.html
Figure 3. Jens Spahn https://www.jens-spahn.de/hubfs/Jens_Spahn_08_breit.jpeg

Electronic Health Records

Central storage of healthcare encounters

Why?

(1) Discover novel disease treatments
(2) Improve patients diagnosis
(3) Improve personalized healthcare

Figure 4. Germany eHealth https://www.krankenkassenzentrale.de/wiki/incoming-en#

Problem Statement

Find a suitable target group of people for a clinical trial test of novel Diabetes drugs or treatments

7 Million cases of **Diabetes** in Germany alone [1]

Figure 7. Diabetes https://www.vectorstock.com/royalty-free-vector/concept-of-diabetes-sugar-vector-18519211
Figure 8. Population https://freesvg.org/vector-image-of-population-icon
[1] https://www.diabetesde.org/ueber_diabetes/was_ist_diabetes_/diabetes_in_zahlen

Related Work

Overview of artificial intelligence in medicine

Amisha[1], Paras Malik[1], Monika Pathania[1], Vyas Kumar Rathaur[2]

[1]Department of Medicine, All India Institute of Medical Sciences (AIIMS), Rishikesh, Uttarakhand, [2]Department of Paediatrics, Government Doon Medical College, Dehradun, Uttarakhand, India

Evaluating predictive modeling algorithms to assess patient eligibility for clinical trials from routine data

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG MEDIZINISCHE FAKULTÄT

Felix Köpcke[1*], Dorota Lubgan[2], Rainer Fietkau[2], Axel Scholler[3], Carla Nau[3], Michael Stürzl[4], Roland Croner[5], Hans-Ulrich Prokosch[1] and Dennis Toddenroth[1]

Artificial Intelligence Applications in Type 2 Diabetes Mellitus Care: Focus on Machine Learning Methods

Shahabeddin Abhari[1], Sharareh R. Niakan Kalhori[1], Mehdi Ebrahimi[2,3], Hajar Hasannejadasl[1], Ali Garavand[4]

[2] AI in Medicine https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6691444/ 2019
[3] Patient Selection https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4029400/pdf/1472-6947-13-134.pdf 2013
[4] AI against Diabetes https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6859270/pdf/hir-25-248.pdf 2020

Research Question

Can Machine Learning be safely applied in the real clinical environment if it just provides enough explainability for its predictions?

(1) How many metrics are enough?
(2) Does a combination of metrics improve the overall explainability?

Figure 9. EKG https://wtig.org/wissensdatenbank/ki/explainable-artificial-intelligence-xai-wenn-sich-ki-selbst-erklaeren-muss/

Methods

(1) Model a Neural Network with TensorFlow

(2) Explain the model (***why***)
with Uncertainty Estimation and Metrics

(3) Explain the predictions (***how***)
with SHAP and LIME

Dataset

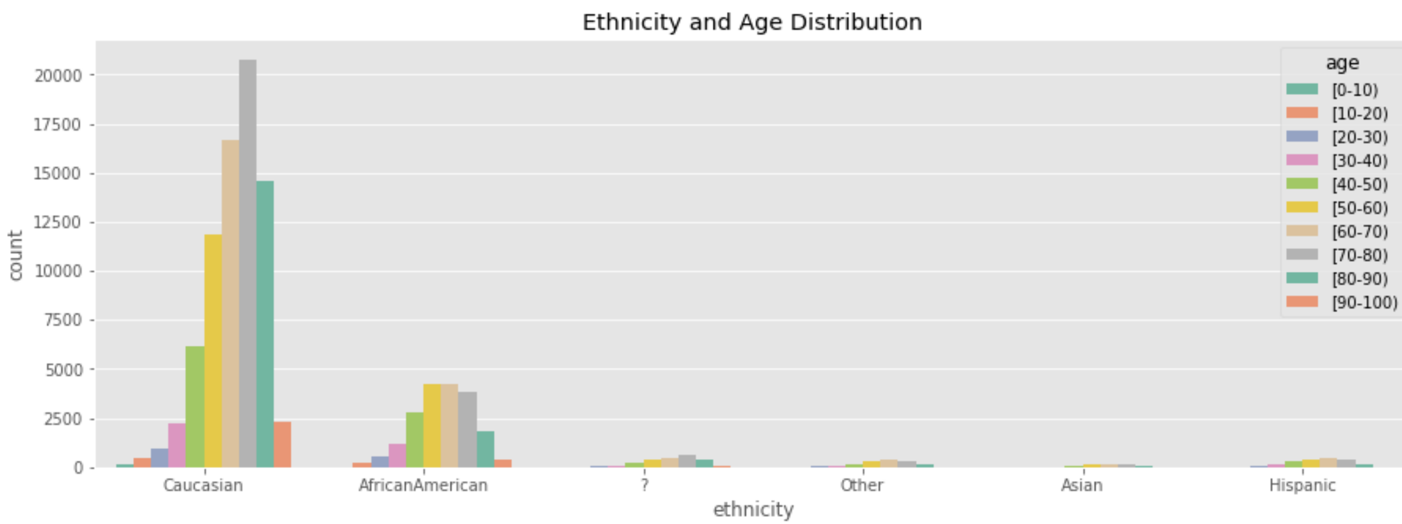
From UCI Machine Learning Repository [5]
- Collected 1999-2008 in the USA
- Originally 101766 Samples and 50 Features

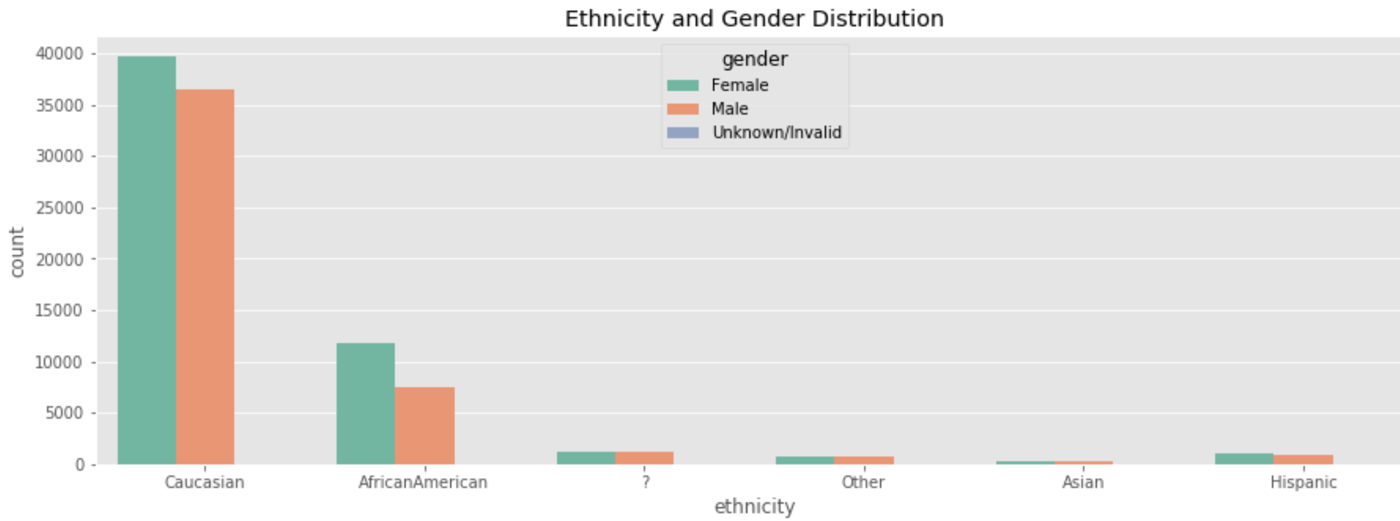
Medication codes from the official FDA website [6]:
- Different company, same drug
  where NDC Code: 71619-388-60 = Glipzide
  and NDC Code: 71619-388-73 = Glipzide

Diagnosis codes from the official ICD website [7]:
- 250.83 stands for type 1 Diabetes

[5] https://archive.ics.uci.edu/ml/datasets/diabetes
[6] https://www.accessdata.fda.gov
[7] https://icdcodelookup.com/icd-10/

Dataset Distribution





Evaluation – Model

(1) Aleatoric Uncertainty
natural
(2) Epistemic Uncertainty
provide more features

Metrics

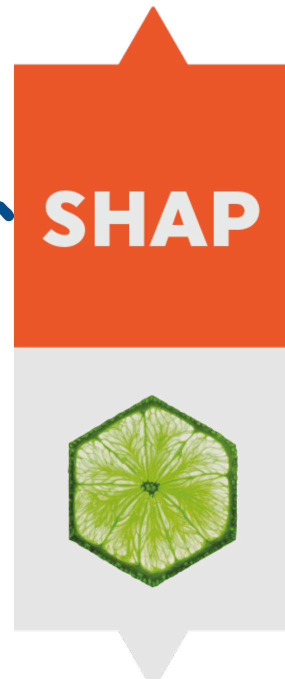
(1) Brier score
(2) AUC score
(3) F1 score
(4) Precision score
(5) Recall Score

Figure 10. Metric SVG https://www.flaticon.com/free-icon/statistics_2004886?term=metrics&page=1&position=24

Evaluation – Predictions

Model Agnostic!

Shapely



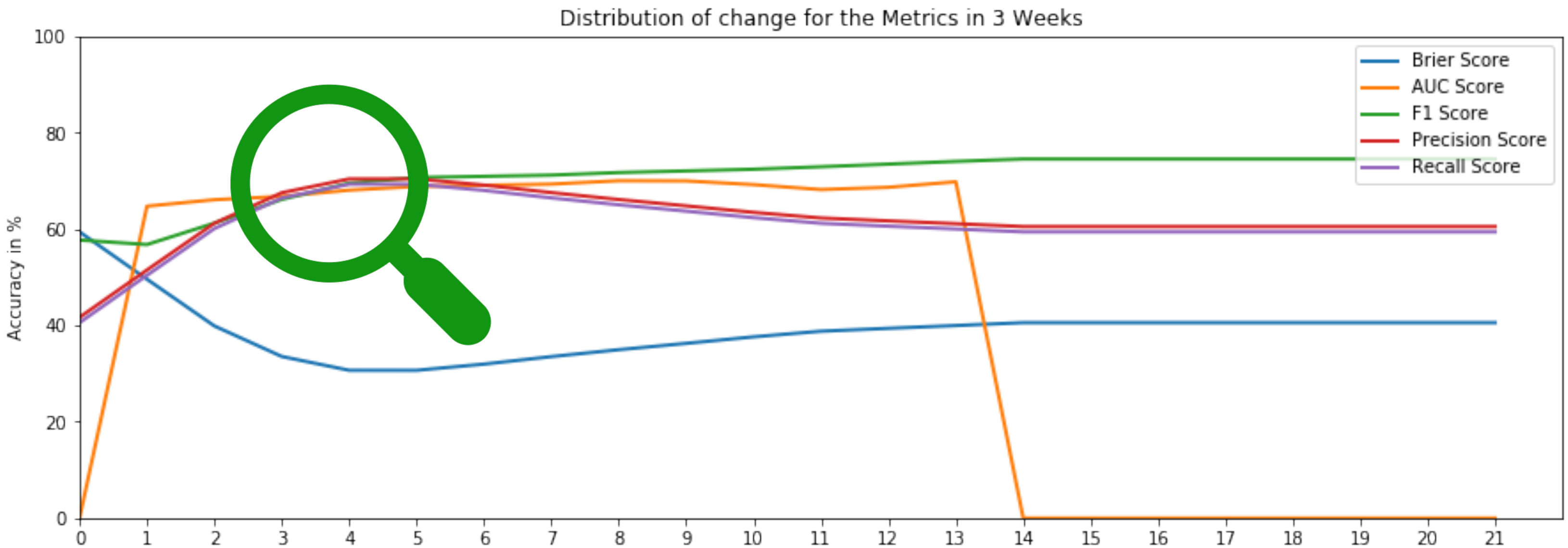
How I got the scores

Most important attributes

LIME

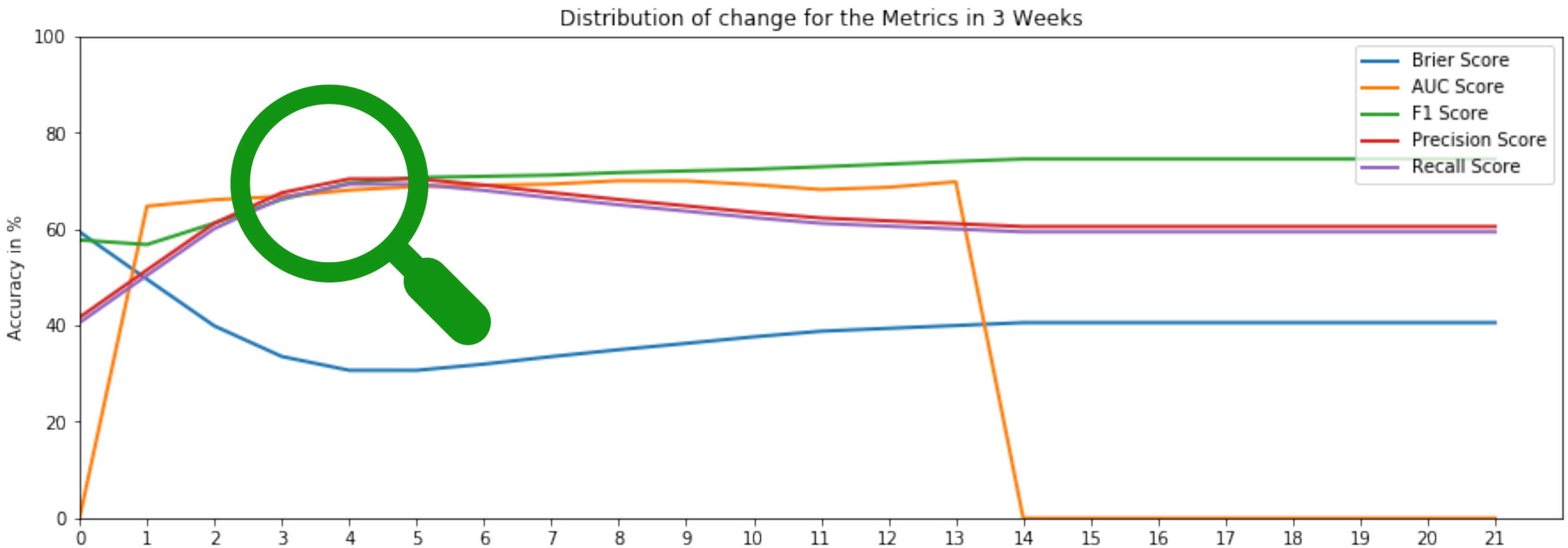
Figure 11. XAI https://codete.com/blog/explainable-ai/

Results – Model Accuracy



(1) Brier Score:	*0.335*
(2) AUC Score:	*0.667*
(3) F1 Score:	*0.661*
(4) Precision Score	*0.664*
(5) Recall Score:	*0.664*

Results – Model Accuracy

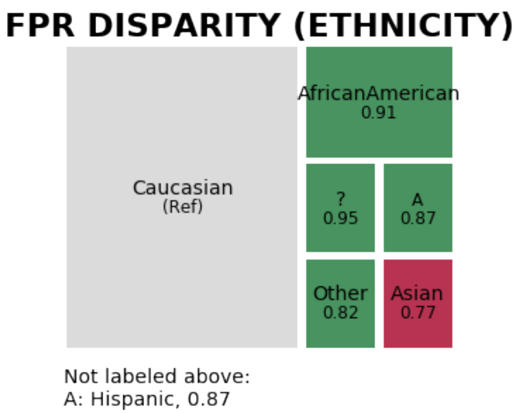


(1) Brier Score: *0.594*
(2) AUC Score: *0.635*
(3) F1 Score: *0.577*
(4) Precision Score *0.405*
(5) Recall Score: *0.405*

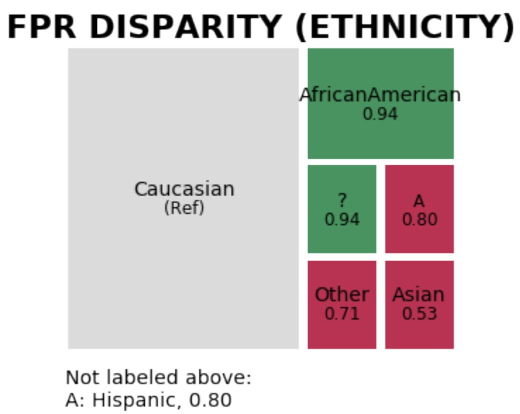
Results – With different boundaries

Ethnicity FPR

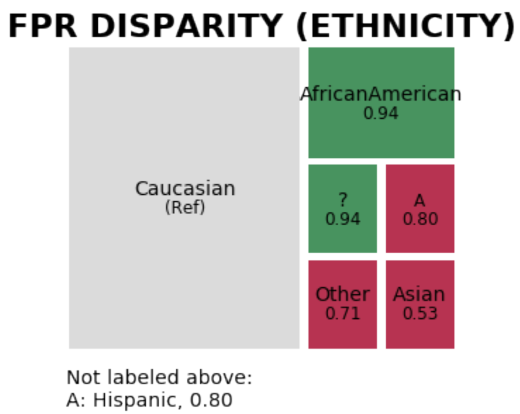
1 week



5 days

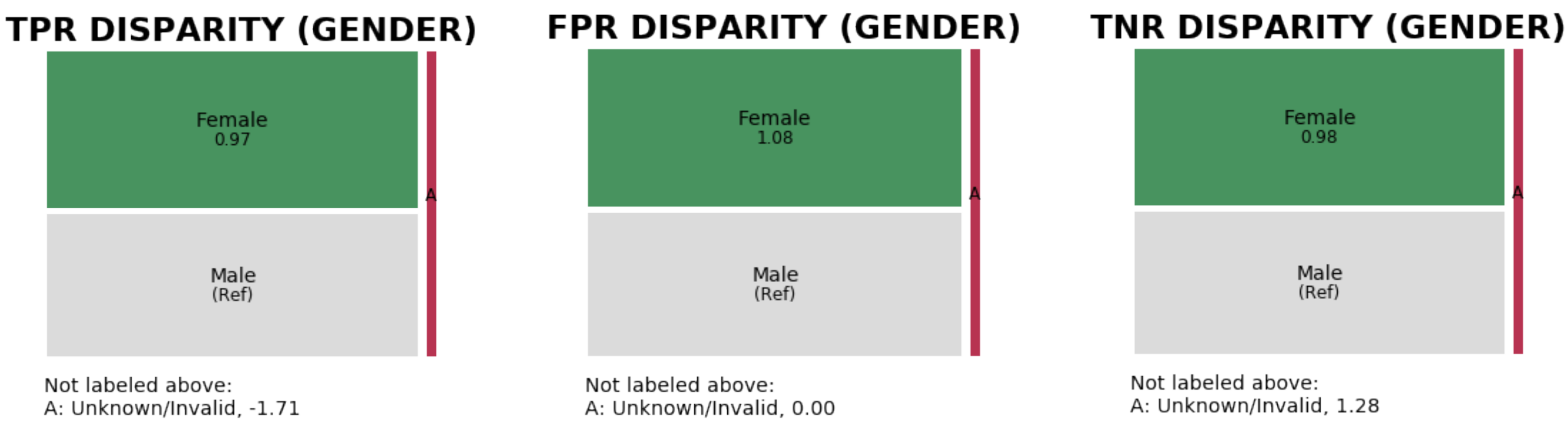


3 days



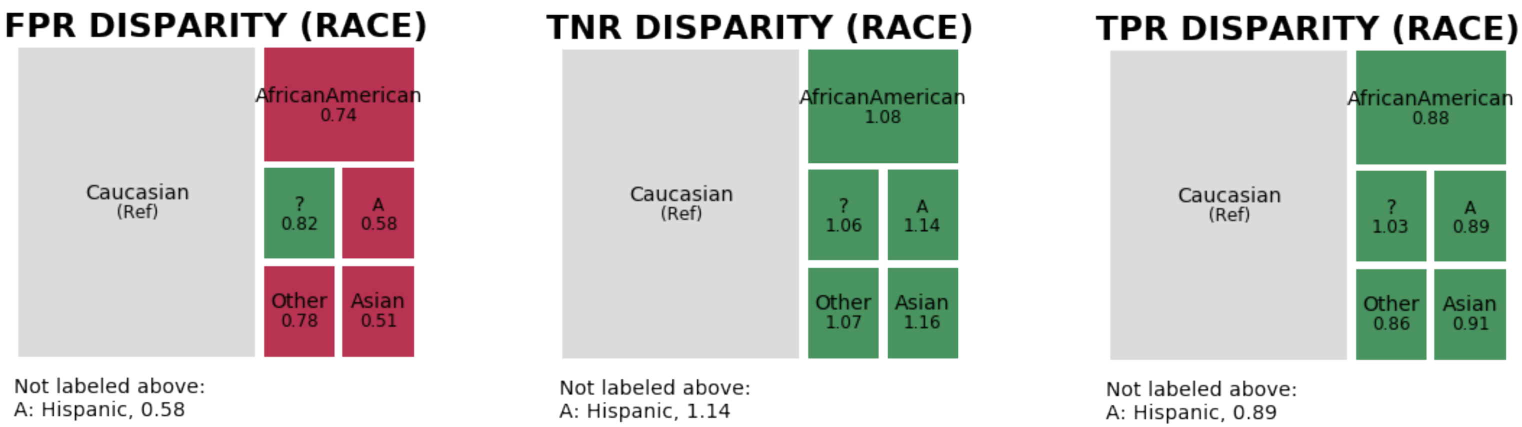
Results – Uncertainty Estimation

Gender Disparity

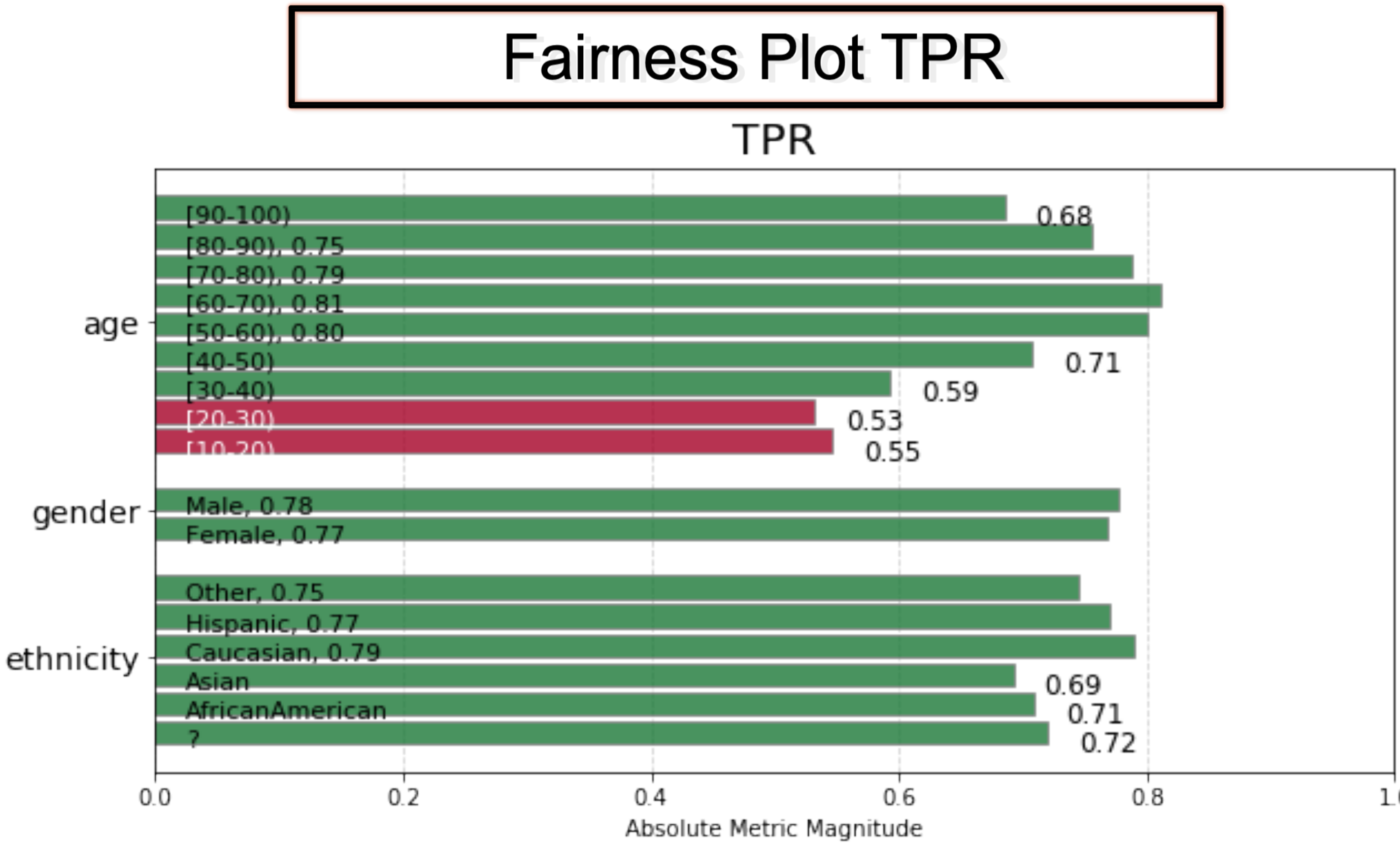


Results – Uncertainty Estimation

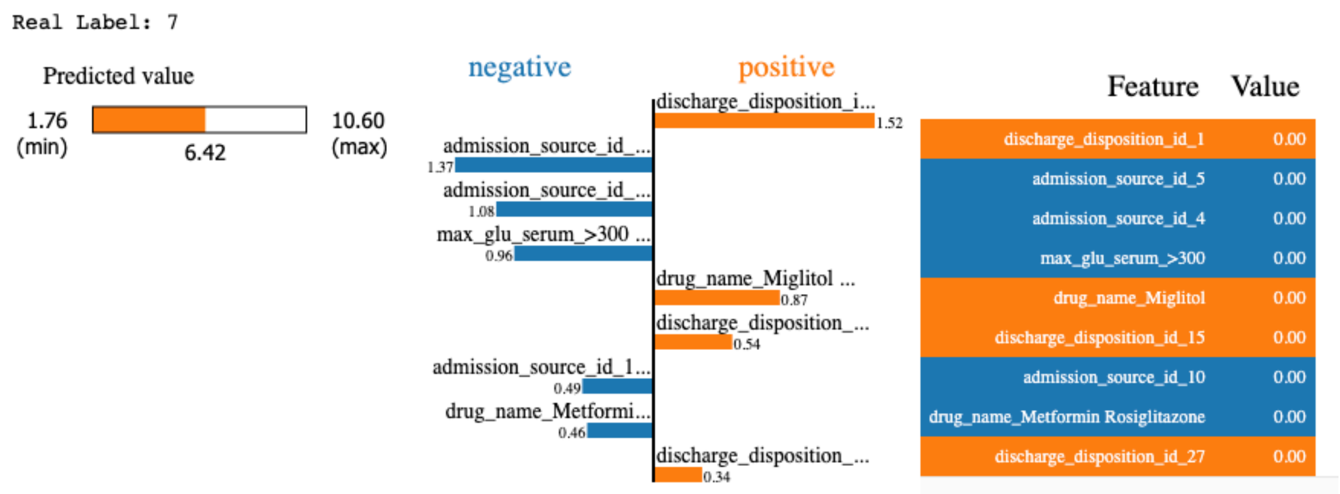
Ethnicity Disparity



Results – Uncertainty Estimation



Results – LIME



Discharge Disposition ID

- discharge_disposition_id_1: Discharged to home
- discharge_disposition_id_15: 'Discharged/transferred within this institution to Medicare approved swing bed'
- discharge_disposition_id_27: NaN
- **discharge_disposition_id_6: Discharged/transferred to home with home health service**

Admission Source ID

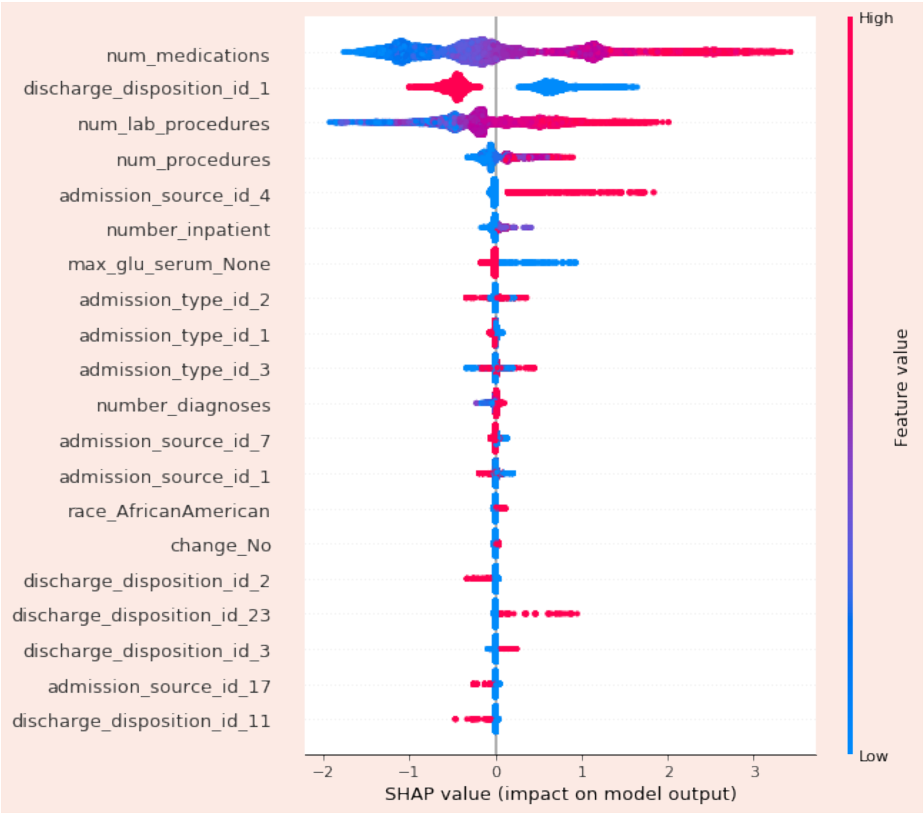
- admission_source_id_5: Transfer from a Skilled Nursing Facility (SNF)
- admission_source_id_4: Transfer from a hospital
- admission_source_id_10: Transfer from critial access hospital
- **admission_source_id_17: NaN**

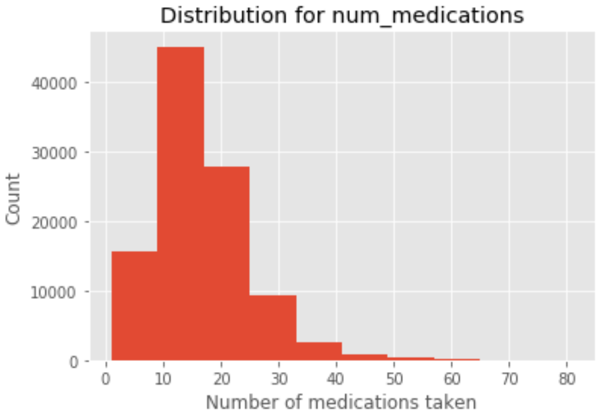
Admission Type ID

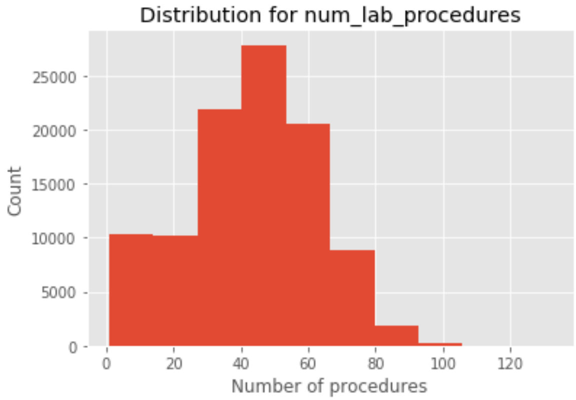
- **admission_type_id_1: Emergency**

race_Caucasian	1
gender_Female	1
age_[70-80)	1
admission_type_id_1	1
discharge_disposition_id_6	1
admission_source_id_17	1
max_glu_serum_>300	1
change_Ch	1
readmitted_NO	1
drug_name_Rosiglitazone	1

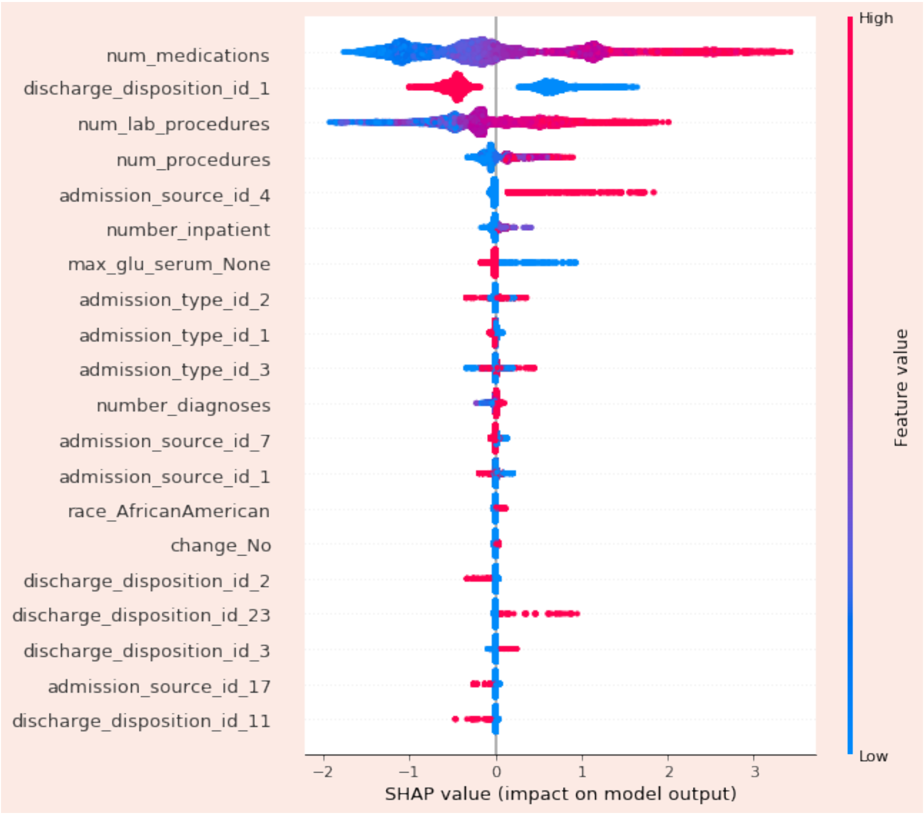
Results – Shapely

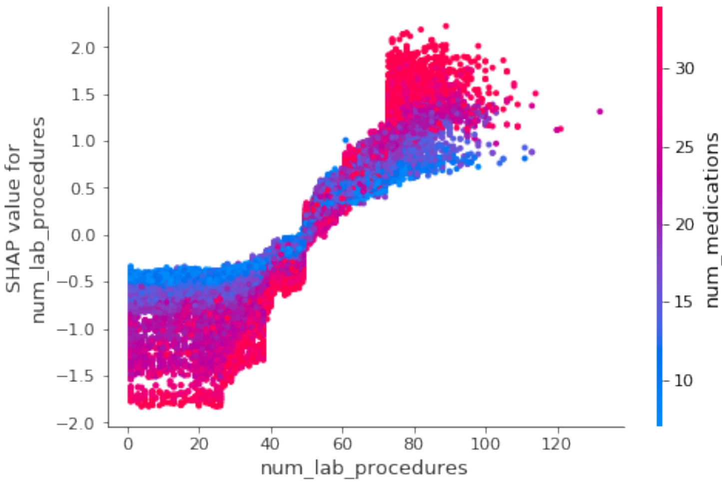


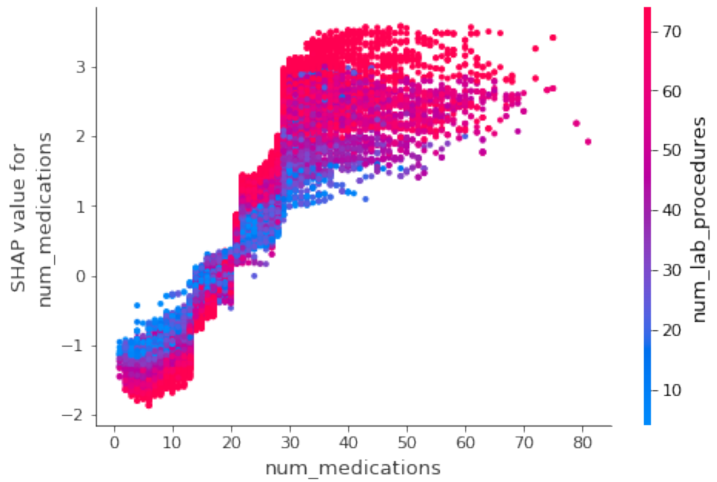




Results – Shapely







Discussion

With just enough explainability,
can Machine Learning be applied in the Industry 4.0?

Pros:

(1) Support for doctor's decision

(2) Patient selection saves a lot of time

(3) Possible new insights

Cons:

(1) Unstructured data

(2) Can it also be applied for other EHR data?

(3) Necessary needed data amount unclear

Conclusion

(1) Apply it without additional doctor’s approval?

For sure not. What if only one person dies because of a false decision of the Machine Learning model?

(2) Apply it as a support technology?

Sure if a doctor validates the result it can be used to support the clinical trials.

Future Work

(1) Try different XAI methods

(2) Try different classifiers:
Use GPU to train XGBoost -> better results?

(3) Epistemic Uncertainty:
Different Neural Network Layers:
tfp.layers.DenseVariational

Thanks for listening