Building meaningful machine learning models for disease prediction

Dr. Shirin Glander

Dep. of Genetic Epidemiology
Institute of Human Genetics
University of Münster

shirin.glander@wwu.de

https://shiring.github.io
https://github.com/ShirinG

31.03.2017

Table of contents

Machine Learning (ML) in disease modeling

⋆

⋆"Machine Learning for Detection and Diagnosis of Disease" (2006). In: *Annual Review of Biomedical Engineering* 8.1. PMID: 16834566, S. 537–565.

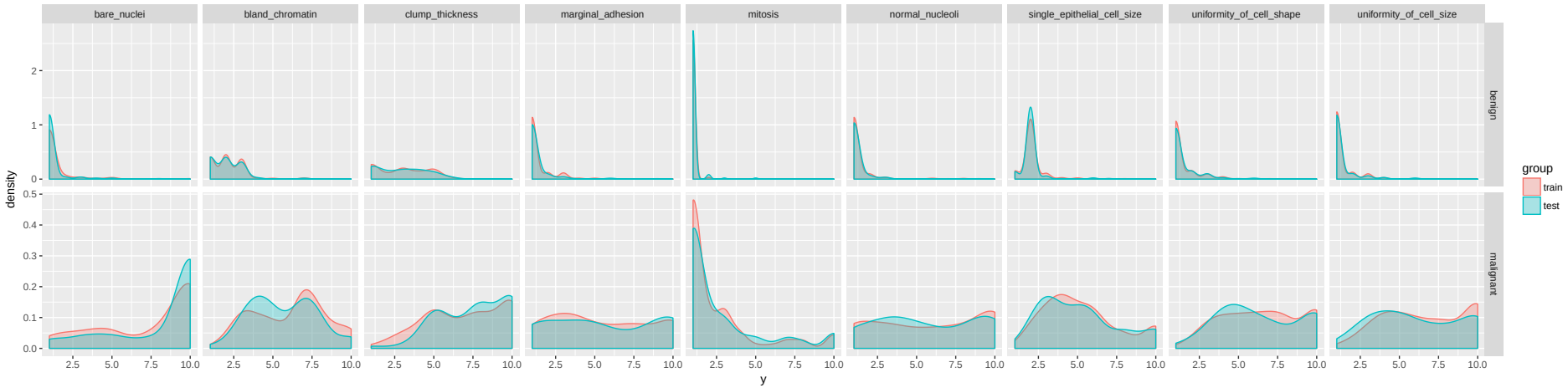
A quick recap of ML basics

How to build ML models in R

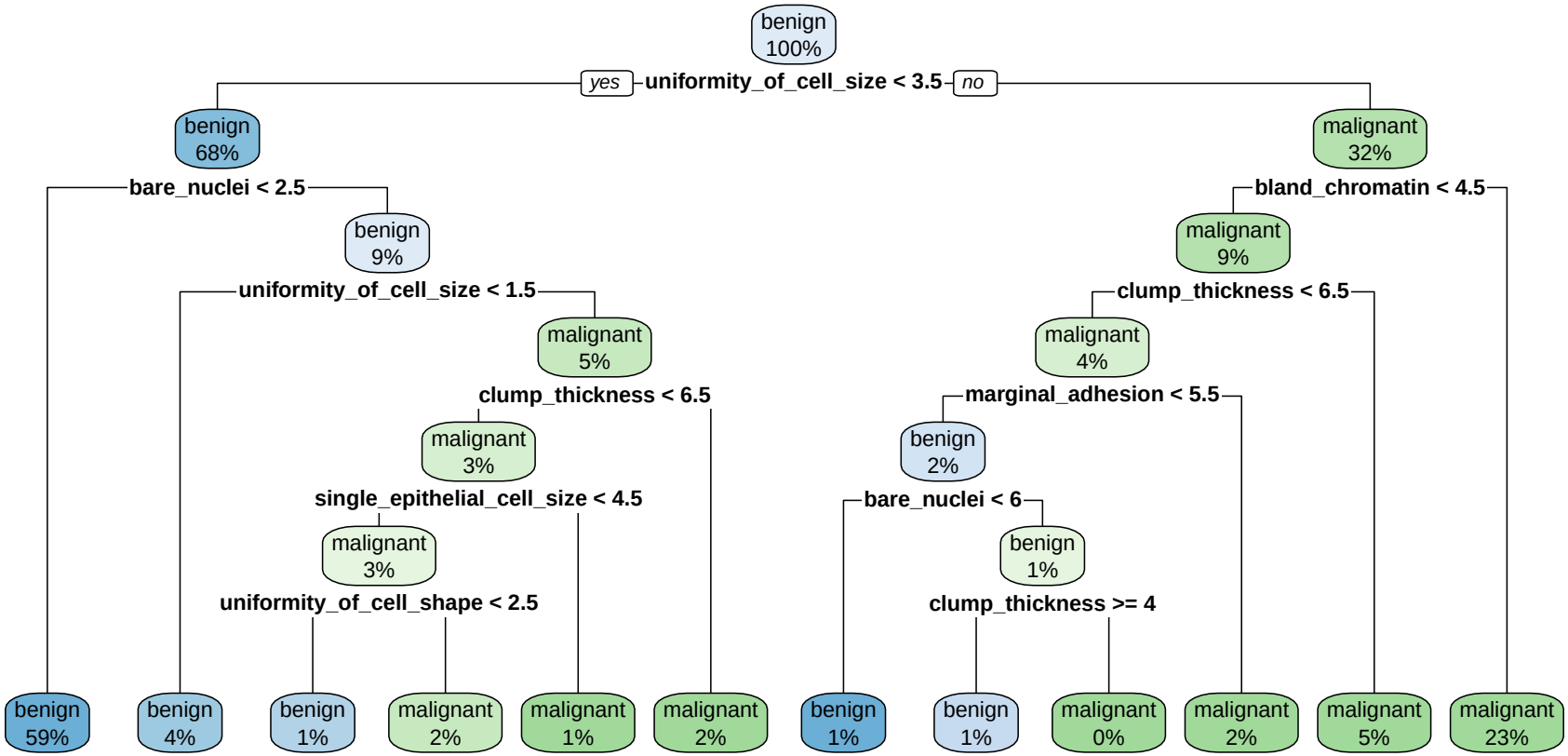
Session setup

Code will be available on my website and on Github
Breast cancer Wisconsin dataset
caret package
h2o package

Distribution

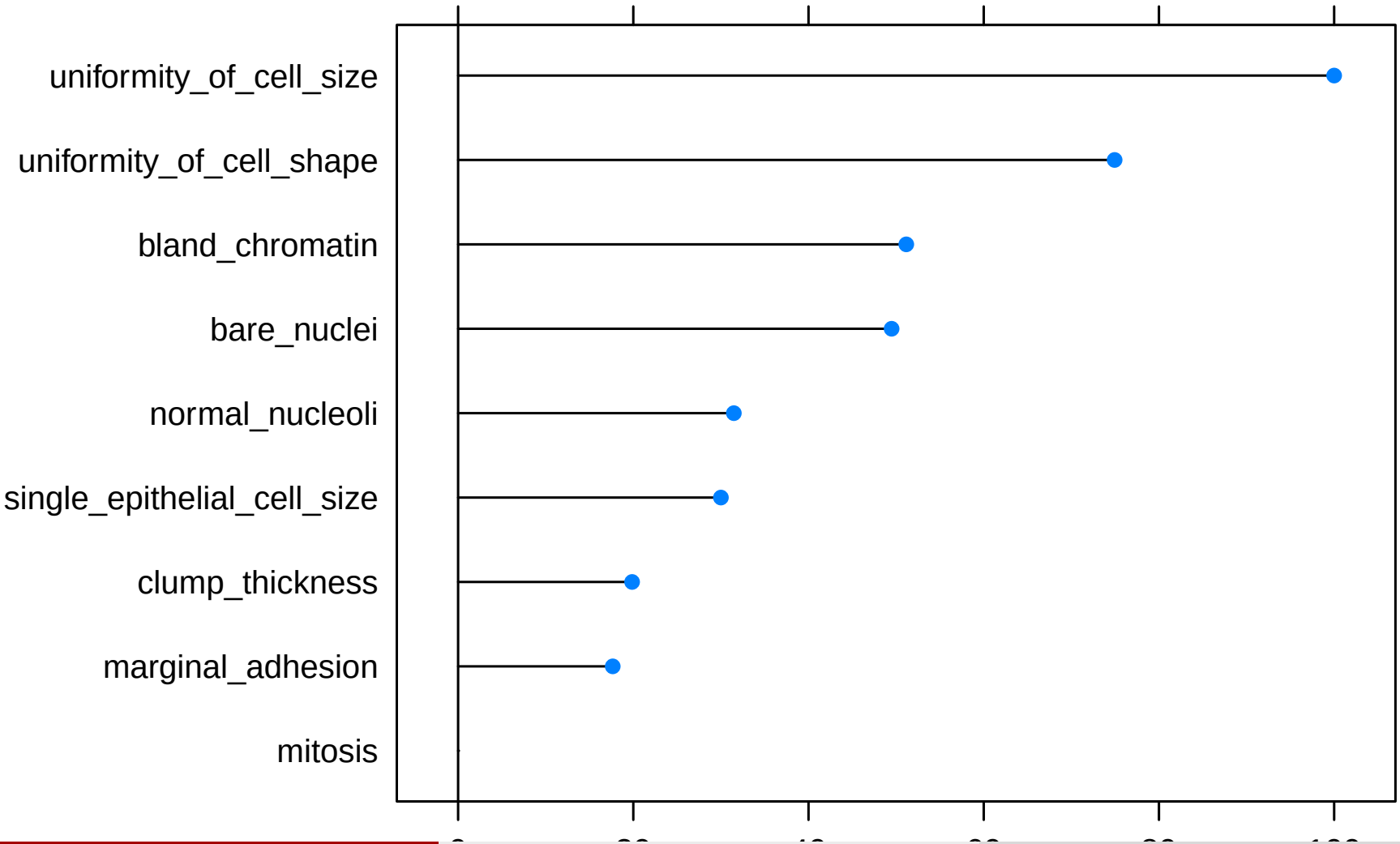


Decision trees

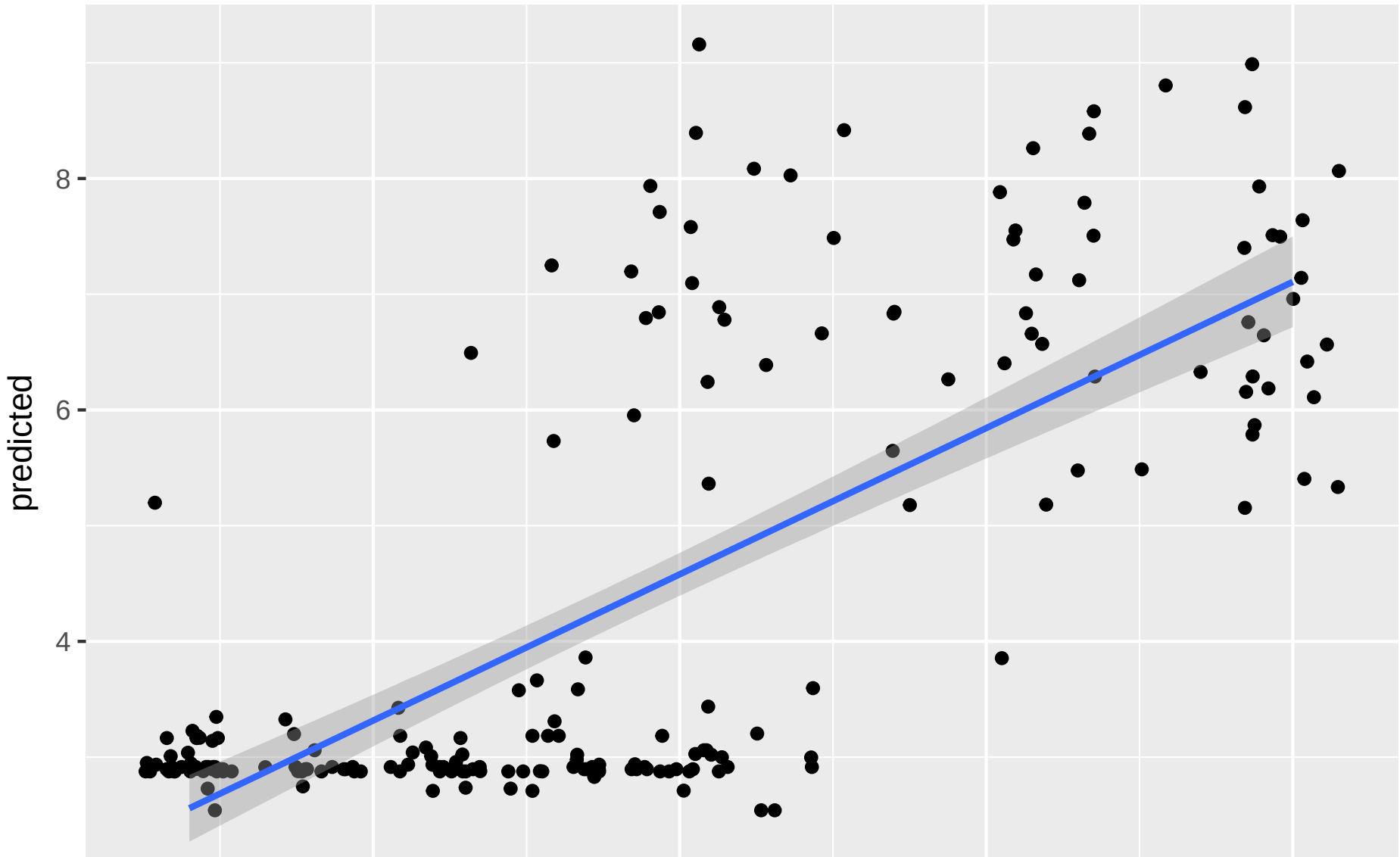


Random Forest

Feature importance

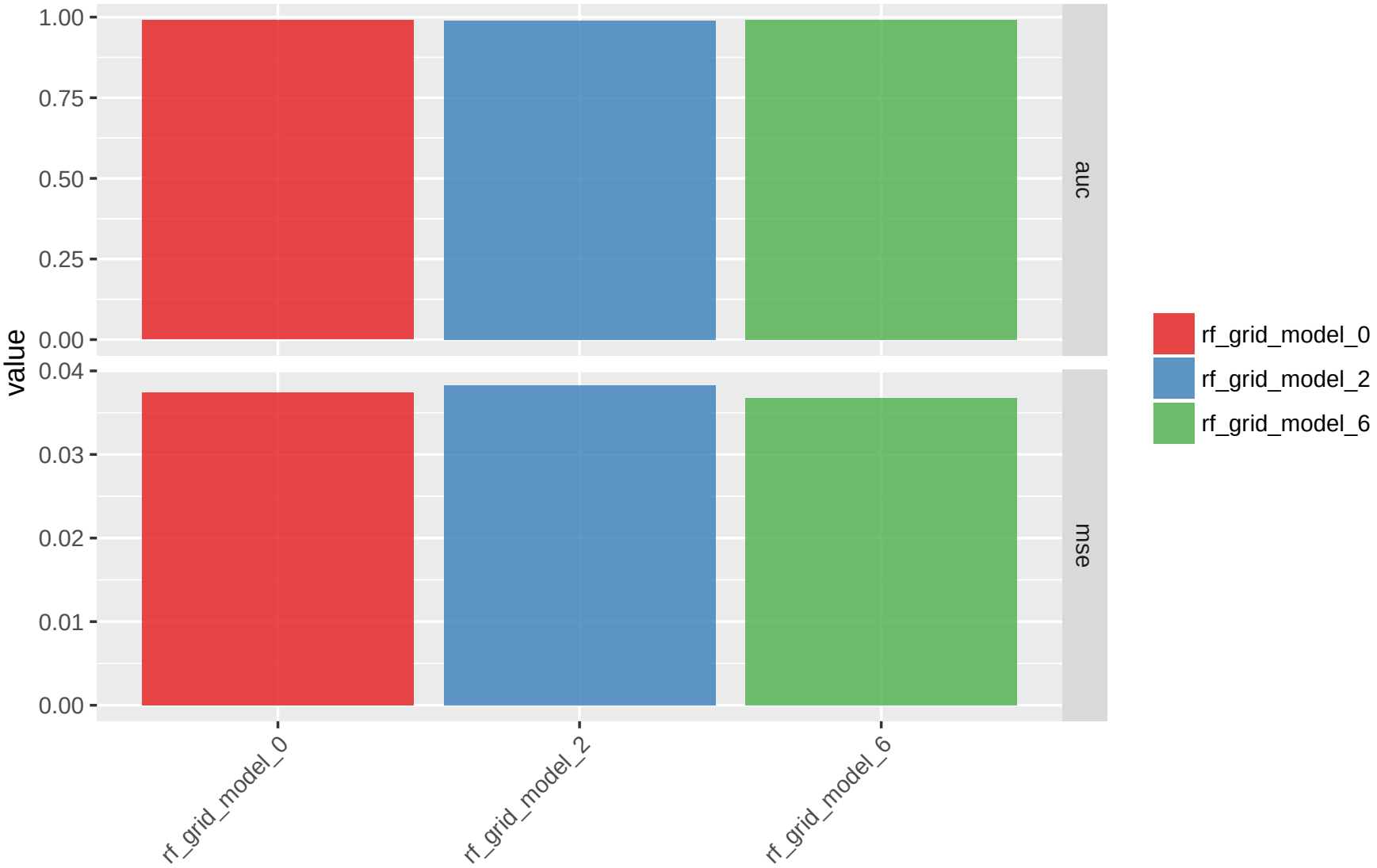


(Generalized) Linear Models

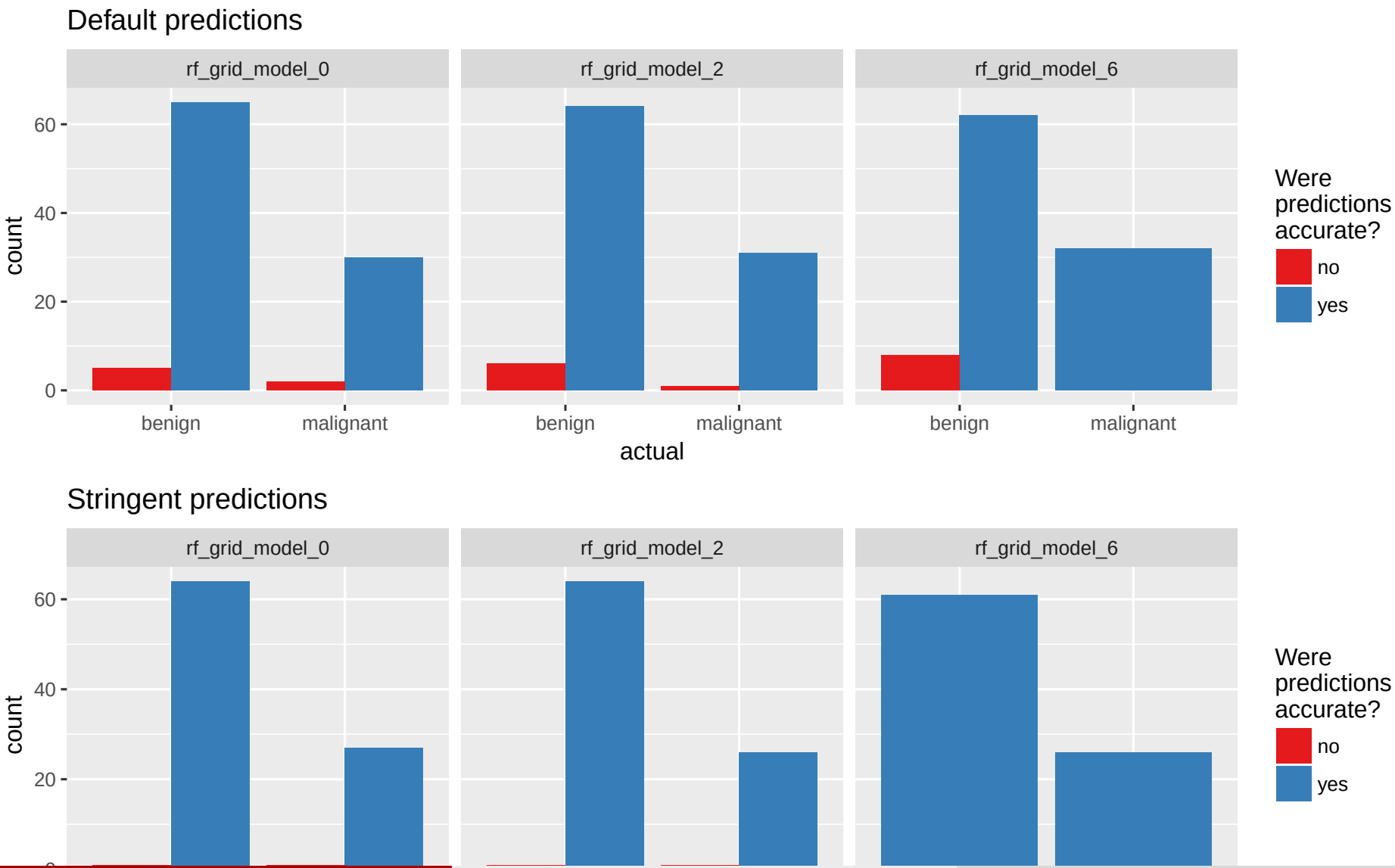


Evaluating ML model performance

AUC and MSE


1.00
0.75
0.50
0.25
0.00
0.04
0.03
0.02
0.01
0.00
value
auc
mse
rf_grid_model_0
rf_grid_model_2
rf_grid_model_6
rf_grid_model_0
rf_grid_model_2
rf_grid_model_6

Predictions on test data



Thank you for your attention!

Questions?

Slides and code will be available on Github:
https://github.com/ShirinG/Webinar_ML_for_disease

Code will also be on my website: https://shiring.github.io

shirin.glander@wwu.de