

# Readme

## Interaction difference test for prediction models

### Online Supplemental Source R-Code

This file contains all source code to reproduce the simulations and data analysis of the manuscript „Interaction difference test for prediction models“. Note that in the simulations many temporary files and folders were created. Therefore it is advised to enable the necessary file system permissions in the root of the simulation folder.

File **Functions.R** includes all necessary definitions (libraries, functions) needed to start any simulation. The general structure of Code includes a **summary** file for each of the three simulations (Article Sections 4.1, 4.2 and 4.3). The summary files include setups to configure simulations and creating required objects. Then the other files are **templates** for specific simulation settings. These were computed in a cluster environment using different design parameters. In particular the code has the following structure:

#### 1. Test statistic distribution in linear models (Article Section 4.1)

##### 1.1. Summary for computation (Article Figure 1, 2):

**testData\_Sim\_NumberCovariatesStatDist\_summary.R**

##### 1.2. Templates:

1.2.1. **testData\_Sim\_NumberCovariatesStatDist\_nullH.R**: This file simulates the test statistic  $z_4$  under  $H_0$  (Article Figure 1)

1.2.2. **testData\_Sim\_NumberCovariatesStatDist\_oneH.R**: This file simulates the test statistic  $z_4$  under  $H_1$  (Article Figure 2)

##### 1.3. Summary for computation of t-statistic (Article Figure 3):

**Sim\_NumberCovariatesStatDist\_oneH\_tValue\_summary.R**

##### 1.4. Template for computation of t-statistic (Article Figure 3):

**Sim\_NumberCovariatesStatDist\_oneH\_tValue.R**

#### 2. Power simulation in linear models (Article Section 4.2)

##### 2.1. Summary: **testData\_Sim\_PowerType1error\_summary.R**

##### 2.2. Templates:

2.2.1. **testData\_Sim\_PowerType1error.R**: This file computes the majority of design scenarios to estimate Power and Type I error in linear models.

2.2.2. **testData\_Sim\_PowerType1error\_extra.R**: This file computes the extra design scenarios to estimate Power and Type I error in linear models.

#### 3. Power simulation on nonlinear models (Article Section 4.3)

##### 3.1. Summary: **testData\_SimData\_summary.R**

3.2. Template: **testData\_SimData.R**

3.3. Other: **dataset\_31\_credit-g.arff**: Data export of credit data from the machine learning repository *OpenML-CC18*.

4. Simulation of test statistic with response information (Article Section 4.4.)

Definitions: **Functions\_Response.R**

Summary: **Sim\_Response.R**

5. Data Analysis example (Article Section 5)

Summary: **DataExample.R**

6. Software: R version 4.2.2 with attached packages

car\_3.1-1  
carData\_3.0-5  
caret\_6.0-93  
combinat\_0.0-8  
farff\_1.1.1  
ggplot2\_3.4.0  
gmp\_0.6-10  
gridExtra\_2.3  
lattice\_0.20-45  
Matrix\_1.5-1  
matrixcalc\_1.0-6  
mvnfast\_0.2.7  
PropCIs\_0.3-0  
ranger\_0.14.1  
readr\_2.1.3  
Rmpfr\_0.9-0  
scales\_1.2.1  
xtable\_1.8-4  
zip\_2.2.2