

# Package ‘DIFboost’

July 22, 2025

**Type** Package

**Title** Detection of Differential Item Functioning (DIF) in Rasch Models  
by Boosting Techniques

**Version** 0.4

**Date** 2025-07-22

**Imports** mboost, penalized, stabs

**Maintainer** Gunther Schaubberger <gunther.schaubberger@tum.de>

**Description** Performs detection of Differential Item Functioning using the method DIFboost as proposed by Schaubberger and Tutz (2016) <[doi:10.1111/bmsp.12060](https://doi.org/10.1111/bmsp.12060)>.

**License** GPL-2

**LazyLoad** yes

**NeedsCompilation** no

**Author** Gunther Schaubberger [aut, cre]

**Repository** CRAN

**Date/Publication** 2025-07-22 11:50:15 UTC

## Contents

|                            |          |
|----------------------------|----------|
| DIFboost-package . . . . . | 2        |
| DIFboost . . . . .         | 3        |
| print.DIFboost . . . . .   | 5        |
| simul.data . . . . .       | 6        |
| <b>Index</b>               | <b>8</b> |

---

DIFboost-package

*DIFboost*

---

## Description

A package to perform DIFboost, a method to detect DIF (Differential Item Functioning) in Rasch Models. It can handle settings with many covariates and also metric covariates simultaneously. The method is described in Tutz and Schauburger (2015). Model/variable selection is performed using stability selection.

## Details

The method assumes the DIFmodel from Tutz and Schauburger (2015) where boosting is used for DIF detection. Computation is based on the functions [gamboost](#) and [stabsel](#).

## Author(s)

Gunther Schauburger  
<[gunther.schauberger@tum.de](mailto:gunther.schauberger@tum.de)>  
<https://www.sg.tum.de/epidemiologie/team/schauberger/>

## References

Schauberger, Gunther and Tutz, Gerhard (2016): *Detection of Differential Item Functioning in Rasch Models by Boosting Techniques*, British Journal of Mathematical and Statistical Psychology, 69(1), 80 - 103

## See Also

[DIFboost](#), [print.DIFboost](#)

## Examples

```
## Not run:
data(simul.data)

Y <- simul.data[,1:10]
X <- simul.data[,11:13]

m1 <- DIFboost(Y = Y, X = X)
print(m1)

## End(Not run)
```

---

|          |                                                                                                |
|----------|------------------------------------------------------------------------------------------------|
| DIFboost | <i>Detection of Differential Item Functioning (DIF) in Rasch Models by Boosting Techniques</i> |
|----------|------------------------------------------------------------------------------------------------|

---

## Description

A function to perform DIFboost, a method to detect DIF (Differential Item Functioning) in Rasch Models. It can handle settings with many covariates and also metric covariates simultaneously. The method is described in Tutz and Schauberger (2015). Model/variable selection is performed using stability selection.

## Usage

```
DIFboost(Y, X, mstop = 400, trace = TRUE, cutoff = 0.9,
         B = 500, mc.cores = 1, q = 0.6 * I)
```

## Arguments

|          |                                                                                                                                                                              |
|----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Y        | Data frame (one row per person, one column per item) containing response. May only contain 0 or 1.                                                                           |
| X        | Data frame (one row per person, one column per covariate) containing covariates. Has to be standardized.                                                                     |
| mstop    | Number of boosting iterations maximally performed in one iteration of the stability selection.                                                                               |
| trace    | Should the trace of the single boosting steps be printed?                                                                                                                    |
| cutoff   | Cutoff value for stability selection.                                                                                                                                        |
| B        | Number of subsamples used for stability selection.                                                                                                                           |
| mc.cores | Number of cores for parallelized stability selection. For windows machines, parallelization is not possible.                                                                 |
| q        | Maximum number of base learner to be included in the boosting algorithm for one subsample in stability selection. By default set to 60 percent of the total number of items. |

## Details

The method assumes the DIFmodel from Tutz and Schauberger (2015) where boosting is used for DIF detection. Computation is based on the functions [gamboost](#) and [stabsel](#).

## Value

|              |                                                                                  |
|--------------|----------------------------------------------------------------------------------|
| model        | Model from initial gamboost fit                                                  |
| dif.mat      | Estimates of the item-specific parameter estimates, with zeros for non-DIF items |
| coefficients | coefficient vector with all estimates from refitted model                        |
| theta        | Estimated person abilities                                                       |

|               |                                                                                                                  |
|---------------|------------------------------------------------------------------------------------------------------------------|
| beta          | Estimated item difficulties                                                                                      |
| gamma         | Estimated item-specific parameters                                                                               |
| P             | Number of (valid) persons                                                                                        |
| I             | Number of items                                                                                                  |
| names.y       | Names of the items                                                                                               |
| names.x       | Names of the covariates                                                                                          |
| design.matrix | Design matrix for refitted model                                                                                 |
| PFER          | upper bound for the per-family error rate. For details see <a href="#">stabsel</a> .                             |
| lin.pred      | linear predictor from refitted model                                                                             |
| DIF.items     | Which items have been detected to be DIF items?                                                                  |
| ref.item      | Reference item                                                                                                   |
| phat          | selection probabilities for single base learners in stability selection. For details see <a href="#">stabsel</a> |
| cutoff        | cutoff value used for stability selection                                                                        |

### Author(s)

Gunther Schaubberger  
 <gunther.schaubberger@tum>  
<https://www.sg.tum.de/epidemiologie/team/schaubberger/>

### References

Schaubberger, Gunther and Tutz, Gerhard (2016): *Detection of Differential Item Functioning in Rasch Models by Boosting Techniques*, British Journal of Mathematical and Statistical Psychology, 69(1), 80 - 103

### See Also

[print.DIFboost](#), [gamboost](#), [stabsel](#)

### Examples

```
## Not run:
data(simul.data)

Y <- simul.data[,1:10]
X <- simul.data[,11:13]

m1 <- DIFboost(Y = Y, X = X)
print(m1)

## End(Not run)
```

---

|                |                                    |
|----------------|------------------------------------|
| print.DIFboost | <i>Print function for DIFboost</i> |
|----------------|------------------------------------|

---

## Description

Prints the most important output of a DIFboost object.

## Usage

```
## S3 method for class 'DIFboost'  
print(x, ...)
```

## Arguments

|     |                                                                       |
|-----|-----------------------------------------------------------------------|
| x   | DIFboost object, created by <a href="#">DIFboost</a>                  |
| ... | Further arguments to be passed to the <a href="#">print</a> function. |

## Author(s)

Gunther Schauburger  
<[gunther.schauburger@tum.de](mailto:gunther.schauburger@tum.de)>  
<https://www.sg.tum.de/epidemiologie/team/schauburger/>

## References

Schauburger, Gunther and Tutz, Gerhard (2016): *Detection of Differential Item Functioning in Rasch Models by Boosting Techniques*, British Journal of Mathematical and Statistical Psychology, 69(1), 80 - 103

## See Also

[DIFboost](#)

## Examples

```
## Not run:  
data(simul.data)  
  
Y <- simul.data[,1:10]  
X <- simul.data[,11:13]  
  
m1 <- DIFboost(Y = Y, X = X)  
print(m1)  
  
## End(Not run)
```

---

simul.data

*Simulated data set*


---

### Description

Simulated data set with 100 persons, 10 items and 3 (standardized) covariates. Items 1, 2 and 3 are DIF items.

### Usage

```
data(simul.data)
```

### Format

```
Item1 Item 1, DIF item
Item2 Item 2, DIF item
Item3 Item 3, DIF item
Item4 Item 4, non-DIF item
Item5 Item 5, non-DIF item
Item6 Item 6, non-DIF item
Item7 Item 7, non-DIF item
Item8 Item 8, non-DIF item
Item9 Item 9, non-DIF item
Item10 Item 10, non-DIF item
CovBin1 Binary covariate (standardized)
CovBin2 Binary covariate (standardized)
CovMet Metric covariate (standardized)
```

### Author(s)

Gunther Schaubberger  
 <gunther.schaubberger@tum>  
<https://www.sg.tum.de/epidemiologie/team/schaubberger/>

### References

Schaubberger, Gunther and Tutz, Gerhard (2016): *Detection of Differential Item Functioning in Rasch Models by Boosting Techniques*, British Journal of Mathematical and Statistical Psychology, 69(1), 80 - 103

### See Also

[DIFboost](#), [print.DIFboost](#)

**Examples**

```
## Not run:  
data(simul.data)  
  
Y <- simul.data[,1:10]  
X <- simul.data[,11:13]  
  
m1 <- DIFboost(Y = Y, X = X)  
print(m1)  
  
## End(Not run)
```

# Index

## \* **DIFboost**

DIFboost, [3](#)

DIFboost, [2](#), [3](#), [5](#), [6](#)

DIFboost-package, [2](#)

gamboost, [2–4](#)

print, [5](#)

print.DIFboost, [2](#), [4](#), [5](#), [6](#)

simul.data, [6](#)

stabsel, [2–4](#)