

Package ‘IFTPredictor’

February 13, 2025

Type Package

Title Predictions Using Item-Focused Tree Models

Version 0.1.0

Author Muditha L. Bodawatte Gedara [aut, cre],
Barret A. Monchka [aut],
Lisa M. Lix [aut]

Maintainer Muditha L. Bodawatte Gedara <muditha.lakmali.1993@gmail.com>

Description This function predicts item response probabilities and item responses using the item-focused tree model. The item-focused tree model combines logistic regression with recursive partitioning to detect Differential Item Functioning in dichotomous items. The model applies partitioning rules to the data, splitting it into homogeneous subgroups, and uses logistic regression within each subgroup to explain the data. Differential Item Functioning detection is achieved by examining potential group differences in item response patterns. This method is useful for understanding how different predictors, such as demographic or psychological factors, influence item responses across subgroups.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Imports DIFtree

Suggests devtools, testthat (>= 3.0.0)

RoxygenNote 7.3.2

Config/testthat/edition 3

NeedsCompilation no

Repository CRAN

Date/Publication 2025-02-13 11:30:28 UTC

Contents

mydata	2
predict_item_responses	3

mydata

Example Dataset for DIFtree

Description

A dataset used for demonstrating the IFTPredictor package. This dataset includes response data (columns 1-20) and covariate data (columns 21-24)

Usage

mydata

Format

A data frame with 500 rows and 24 variables:

- V1** Binary response variable (numeric: 0 or 1) representing item 1.
- V2** Binary response variable (numeric: 0 or 1) representing item 2.
- V3** Binary response variable (numeric: 0 or 1) representing item 3.
- V4** Binary response variable (numeric: 0 or 1) representing item 4.
- V5** Binary response variable (numeric: 0 or 1) representing item 5.
- V6** Binary response variable (numeric: 0 or 1) representing item 6.
- V7** Binary response variable (numeric: 0 or 1) representing item 7.
- V8** Binary response variable (numeric: 0 or 1) representing item 8.
- V9** Binary response variable (numeric: 0 or 1) representing item 9.
- V10** Binary response variable (numeric: 0 or 1) representing item 10.
- V11** Binary response variable (numeric: 0 or 1) representing item 11.
- V12** Binary response variable (numeric: 0 or 1) representing item 12.
- V13** Binary response variable (numeric: 0 or 1) representing item 13.
- V14** Binary response variable (numeric: 0 or 1) representing item 14.
- V15** Binary response variable (numeric: 0 or 1) representing item 15.
- V16** Binary response variable (numeric: 0 or 1) representing item 16.
- V17** Binary response variable (numeric: 0 or 1) representing item 17.
- V18** Binary response variable (numeric: 0 or 1) representing item 18.
- V19** Binary response variable (numeric: 0 or 1) representing item 19.
- V20** Binary response variable (numeric: 0 or 1) representing item 20.
- x1** An integer variable (e.g., a grouping factor or categorical covariate).
- x2** A continuous numeric covariate.
- x3** An integer variable (binary: 0 or 1).
- x4** A continuous numeric covariate.

Examples

```
data("mydata", package = "IFTPredictor")
head(mydata)
```

```
predict_item_responses
```

Predictions Using Item-Focused Tree Models

Description

This function predicts item response probabilities and item responses using the item-focused tree (IFT) model. The item-focused tree model combines logistic regression with recursive partitioning to detect Differential Item Functioning (DIF) in dichotomous items. The model applies partitioning rules to the data, splitting it into homogeneous subgroups, and uses logistic regression within each subgroup to explain the data. Differential Item Functioning detection is achieved by examining potential group differences in item response patterns. This model produces tree diagrams to visualize homogeneous subgroups within the population exhibiting similar response patterns and may therefore be helpful for developing personalized interventions and optimizing resource allocation in healthcare.

Usage

```
predict_item_responses(model, dataset, total_score)
```

Arguments

model	A DIFtree model object.
dataset	A data frame containing the required data. The 'total_score' column must be included.
total_score	The name of the column in the dataset representing the total score (e.g., "total_score").

Details

The logistic regression model for the i -th PROM item is defined as:

$$\log \left(\frac{P(Y_{pi} = 1 | S_p, g)}{P(Y_{pi} = 0 | S_p, g)} \right) = \eta_{pi} = \beta_{0i} + S_p \beta_i + \gamma_{ig},$$

where, $Y_{pi} \in \{0, 1\}$: The response of person p to the i -th item. $p = 1, 2, \dots, P$: The number of persons. $i = 1, 2, \dots, I$: The number of items. g : Group membership ($g = 0$ for the reference group, $g = 1$ for the focal group). S_p : The ability level (e.g., total PROM score) of person p . β_{0i} : The intercept or item difficulty parameter. β_i : The slope or item discrimination parameter. γ_{ig} : The group-specific parameter.

The IFT model extends this logistic regression model for DIF detection for the i -th PROM item:

$$\eta_{pi} = \beta_i S_p + [\gamma_{ik} I(x_{pl} \leq c_l) + \gamma_{ir} I(x_{pl} > c_l)],$$

where, $l = 1, \dots, L$: The number of partitions. c_l : The threshold for the l -th variable. $x_{pl} \leq c_l$ and $x_{pl} > c_l$: The subgroups defined by tree partitions. $I(\cdot)$: The indicator function (1 if true, 0 otherwise). γ_{ik} and γ_{ir} : Subgroup-specific intercepts for logistic regression models in partitioned regions. The terminal or leaf nodes of the tree represent the final groups of patients with similar response patterns. The IFT model assumes unidimensionality, and the covariates x_{pl} can be either continuous or categorical.

If an item is never chosen for splitting, it is assumed to be free of DIF. The equation for an item free of DIF can be defined as:

$$\eta_{pi} = \beta_i S_p + \beta_0 i,$$

Value

A list containing:

equations	A set of logistic regression equations generated for each item.
predictions	A dataset with predicted probabilities (p) and item responses (I), where $I = 1$ if $p \geq 0.5$, and $I = 0$ otherwise.

Author(s)

Muditha L. Bodawatte Gedara (muditha.lakmali.1993@gmail.com), Barret A. Monchka, Lisa M. Lix

References

Berger, Moritz and Tutz, Gerhard (2016): Detection of Uniform and Non-Uniform Differential Item Functioning by Item Focused Trees, *Journal of Educational and Behavioral Statistics* 41(6), 559-592.

See Also

[DIFtree](#) for training the DIFtree model.

Examples

```
if (requireNamespace("DIFtree", quietly = TRUE)) {
  # Load DIFtree
  library(DIFtree)

  # Load the dataset
  data("mydata", package = "IFTPredictor")

  # Observe the data
  head(mydata)

  # Extract response and covariate data
  mydata<-mydata[,c(4:6,21:23)] # Items 4 to 6 will be tested for DIF across covariates 21 to 23
  Y <- mydata[, 1:3] # Specify items you want to test for DIF (i.e., items 1 to 3)
```

```
X <- mydata[, 4:6] # Covariates

# Create total score column calculating total item score for each patient
mydata$total_score <- rowSums(mydata[, 1:3])

# Fit the DIFtree model (Y = response data, X = covariate data)
mod <- DIFtree(Y, X, model = "Logistic", type = "udif", alpha = 0.05, nperm = 5, trace = TRUE)

# Predict item responses using the model and the total score
result <- predict_item_responses(mod, dataset = mydata, total_score = "total_score")

} else {
  message("The 'DIFtree' package is not installed. Please install it to run this example.")
}
```

Index

* **datasets**

mydata, [2](#)

DIFtree, [4](#)

mydata, [2](#)

predict_item_responses, [3](#)