

Package ‘exstatlink’

December 13, 2022

Title To find the statistical relationships between exposure factors
and health outcome.

Version 1.0.0

Description To find the statistical relationships between exposure factors and health outcome.

License GPL (>= 3)

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.2

Imports htr,vroom,ggplot2,readxl,gridExtra

NeedsCompilation no

Author bin Wang [aut, cph, cre]

Maintainer bin Wang <binwang@pku.edu.cn>

R topics documented:

| | |
|------------------------|----------|
| FuncExit | 1 |
| InitStatLink | 2 |
| LoadStatLink | 3 |
| StatLinkCros | 3 |
| Index | 6 |

| | |
|----------|--------------------------------|
| FuncExit | <i>End the module analysis</i> |
|----------|--------------------------------|

Description

End the module analysis

Usage

FuncExit(PID)

Arguments

PID chr. Program ID. It must be the same with the PID generated by any initial functions.

Value

Exit status

Author(s)

Bin Wang (corresponding author)

Examples

```
res <- InitStatLink()
res1 = LoadStatLink(PID = res$PID, UseExample = "example#1")
FuncExit(PID = res$PID)
```

InitStatLink

Initialize ExpoStatLink module

Description

Initialize ExpoStatLink module analysis. It can generate an R6 class object.

Usage

```
InitStatLink()
```

Details

ExpoStatLink module is designed to find the statistical relationships between exposure factors and health outcome.

Value

An R6 class object.

Author(s)

Bin Wang

Examples

```
res <- InitStatLink()
```

| | |
|--------------|---|
| LoadStatLink | <i>Load data file for ExpoStatLink for module</i> |
|--------------|---|

Description

Load data file for ExpoStatLink module

Usage

```
LoadStatLink(PID, UseExample = "default", DataPath=NULL, VocaPath=NULL)
```

Arguments

| | |
|------------|---|
| PID | chr. Program ID. It must be the same with the PID generated by ExpoStatLink |
| UseExample | chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module. |
| DataPath | chr. Input directory of data file, e.g. "D:/test/eg_data_biolink.xlsx". It should be noted that the slash symbol is "/", not "\". |
| VocaPath | chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_biolink.xlsx". It should be noted that the slash symbol is "/", not "\". |

Value

An R6 class object containing the input data.

Author(s)

Bin Wang

Examples

```
res <- InitStatLink()
res = LoadStatLink(PID = res$PID, UseExample = "example#1")
```

| | |
|--------------|---|
| StatLinkCros | <i>Build statistical link for cross-sectional data.</i> |
|--------------|---|

Description

Build statistical link for cross-sectional data.

Usage

```
StatLinkCros(PID, OutPath = "default", VarsY, VarsX, LinkModel = "ranger",
  ObsrPartType = "raw", ObsrPartNum = "50", ObsrProfType = "partial",
  ObsrProfNum = "100", ObsrProfVars = "all.x", ObsrProfGeom = "profiles",
  SubjPredSeq = "none", SubjPartType = "break_down")
```

Arguments

| | |
|--------------|---|
| PID | chr. Program ID. It must be the same with the PID generated by ExpoStatLink |
| OutPath | chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set. |
| VarsY | chr. Outcome variable used for modelling. Only one variable can be entered. |
| VarsX | chr. Exposure variable used for modelling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3" |
| LinkModel | chr. Methods to interpret the model. Options include "ranger" (random forest), "glmnet" (elastic net), "svm" (support vector machine), "glm" (linear regression), "gam" (generalized additive model), and "xgboost" (eXtreme gradient boosting). |
| ObsrPartType | chr. Type of transformation that should be applied for dropout loss. Options include "raw" (drop losses), "ratio" (drop_loss/drop_loss_full_model), and "difference" (drop_loss - drop_loss_full_model) |
| ObsrPartNum | chr. Number of observations that should be sampled for calculation of variable importance. The default means variable importance will be calculated on whole dataset (no sampling). If "default", use all Obsrervations. |
| ObsrProfType | chr. Type of variable profile. Options include "partial", "conditional", and "accumulated" |
| ObsrProfNum | int. Number of observations used for calculation of aggregated profiles. By default 100. |
| ObsrProfVars | chr. Names of variables to be explained. If "all.x", all "X variable" above are chosen. |
| ObsrProfGeom | chr. Layout of the explanation profile in dataset level including "aggregates", "profiles" or "points". |
| SubjPredSeq | chr. Subjects which need explanation. Options include "all" (all the subjects), "none" (no subjects), and "other" (copy the subject list by clicking "Available vars"). |
| SubjPartType | chr. Layout of the explanation profile in subject level. Options include "shap", "oscillations", "break_down", and "none". |

Value

A list containing all the statistical explanation results.

Author(s)

Bin Wang

Examples

```
res <- InitStatLink()
res1 = LoadStatLink(PID = res$PID, UseExample = "example#1")
res2 = StatLinkCros(PID=res$PID, VarsY = "Y1" ,VarsX = "all.x",
LinkModel = "ranger",ObsrPartType = "raw" ,ObsrPartNum = "50",
ObsrProfType = "partial" ,ObsrProfNum = "100", ObsrProfVars = "all.x",
ObsrProfGeom = "profiles",SubjPredSeq = "S1,S2,S3",SubjPartType = "break_down")
```

Index

FuncExit, [1](#)

InitStatLink, [2](#)

LoadStatLink, [3](#)

StatLinkCros, [3](#)