Package ‘IVFPred’

December 7, 2023

Title Predict the IVF-ET Outcome

Version 1.0.0

Description ``IVFPred'' is designed to improve model efficiency in predicting pregnancy success rates in IVF-ET and screening the key influencing factors for the individual centers and patients.

License GPL (>= 3)

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.3

Imports httr,ggplot2,writexl,gridExtra,dplyr,readxl,vroom,tidyverse

NeedsCompilation no

Author Weinan Lin [aut, cre],
Bin Wang [aut, cph]

Maintainer Weinan Lin <weinanlin@pku.edu.cn>

R topics documented:

DelLowVarIVF *Delete variables with low variance for IVF-ET data*

Description

Delete variables with low variance

Usage

```
DelLowVarIVF(PID, OutPath = "default")
```

Arguments

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

Value

An R6 class object containing the variable(s) with acceptable variance.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitIVF()
    res1 = LoadIVF(PID = res$PID, UseExample = "example#1")
    res2 = DelLowVarIVF(PID=res$PID)
    FuncExit(PID = res$PID)
```

DelMissIVF *Delete variables with missing values for IVF-ET data*

Description

Delete missing variables with low variance

Usage

```
DelMissIVF(PID, OutPath = "default")
```

Arguments

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

Value

An R6 class object containing the variable(s) without missing values.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitIVF()
    res1 = LoadIVF(PID = res$PID, UseExample = "example#1")
    res2 = DelMissIVF(PID=res$PID)
    FuncExit(PID = res$PID)
```

FuncExit *End the module analysis*

Description

End the module analysis

Usage

```
FuncExit(PID)
```

Arguments

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

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

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

InitIVF *Initialize IVFPred module*

Description

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

Usage

```
InitIVF()
```

Details

IVFPred module is designed to improve model efficiency in predicting pregnancy success rates in IVF-ET and screening the key influencing factors for the individual centers and patients.

Value

An R6 class object.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitIVF()
```

IVFPred *Build predict model for IVF-ET data*

Description

Build predict model for IVF-ET data

Usage

```
IVFPred(PID, OutPath = "default", AutTuneM = "random_search", AutTuneN = "5", RsmpMethod = "cv", Fol
```

Arguments

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

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.

AutTuneM chr. Method for hyper-parameter autotuning. Options include "default", "random_search", "grid_search", "nloptr"(Non-linear optimization), and "gensa"(Generalized simulated annealing). The "default" option uses the simple training method for parameter optimization of mlr3 package.

`AutTuneN` num. Upper limit of model tuning times. It should be more than 20 times to search the appropriate parameters, but it takes more time. In theory, more time, better training results.

`RsmpMethod` chr. Method for resampling. Options include "cv"(cross validation), "loo"(leave-one-out cross validation), "bootstrap"(bootstrapping), "holdout"(holdout).

`Folds` num. Folds for cross validation resampling method. The default value is 5.

`Ratio` num. Ratio for "Holdout" resampling method. The default value is 5.

`Repeats` num. Repeats for "Bootstrap" resampling method.

`Learners` chr. Learners for build IVF-ET outcome predicting model. Options include "lasso", "elastic net", "rf"(Random forest), "stacked generalization", "logistic" and "xgboost"(Xgboost). One or more arbitrary options can be selected at the same time. Note that separates different learners by "," and without space(e.g. Learners ="lasso,rf,xgboost").

Value

An R6 class object containing the model results

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
  res <- InitIVF()
 res1 = LoadIVF(PID = res$PID, UseExample = "example#2")
 res2 = TransImputIVF(PID=res$PID, Vars="all.x")
 res2 = DelLowVarIVF(PID=res$PID)
 res2 = DelMissIVF(PID=res$PID)
res2 = TransTypeIVF(PID = res$PID, Vars = "Y,X85,X86,X87,X88,X89,X90,X91,X92,X93,X94,X95,X96,X97,X98",
                    To = "factor")
 res2 = TransClassIVF(PID = res$PID, Vars = "X47", LevelTo = 4)
 res2 = TransScaleIVF(PID = res$PID, Vars = "X55", Method = "normal")
 res2 = TransDistrIVF(PID = res$PID, Vars = "X61", Method = "log10")
 res3 = IVFPred(PID = res$PID, AutTuneM = "random_search", AutTuneN = 5,
                RsmpMethod = "cv", Folds = 5,
                Learners = "logistic,rf,xgboost")
 FuncExit(PID = res$PID)
```

`IVFValid` *Validation for IVF-ET data*

Description

Build Vaalidation model for IVF-ET data

Usage

```
IVFValid(PID, OutPath = "default", Learner = 'default', VarsNum = 2, AutTuneM = "random_search", Aut
```

Arguments

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

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.

Learner chr; learners for build model to validate and explain; if "default" best model stored in "IVFPred" will be used

VarsNum num. Number of common features include in all groups;

AutTuneM chr. Method for hyper-parameter autotuning. Options include "default", "random_search", "grid_search", "nloptr"(Non-linear optimization), and "gensa"(Generalized simulated annealing). The "default" option uses the simple training method for parameter optimization of mlr3 package.

AutTuneN num. Upper limit of model tuning times. It should be more than 20 times to search the appropriate parameters, but it takes more time. In theory, more time, better training results.

SingleGroup lgl; Whether only single group is contained

Value

An R6 class object containing the model results

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
   res <- InitIVF()
  res1 = LoadIVF(PID = res$PID, UseExample = "example#2")
  res2 = TransImputIVF(PID=res$PID, Vars="all.x")
  res2 = DelLowVarIVF(PID=res$PID)
  res2 = DelMissIVF(PID=res$PID)
 res2 = TransTypeIVF(PID = res$PID, Vars = "Y,X85,X86,X87,X88,X89,X90,X91,X92,X93,X94,X95,X96,X97,X98",
                     To = "factor")
  res2 = TransClassIVF(PID = res$PID, Vars = "X47", LevelTo = 4)
  res2 = TransScaleIVF(PID = res$PID, Vars = "X55", Method = "normal")
  res2 = TransDistrIVF(PID = res$PID, Vars = "X61", Method = "log10")
  res3 = IVFPred(PID = res$PID, AutTuneM = "random_search", AutTuneN = 5,
                RsmpMethod = "cv", Folds = 5,
                Learners = "logistic,rf,xgboost")
  res4 = IVFValid(PID = res$PID, VarsNum = 2,SingleGroup = TRUE)
  FuncExit(PID = res$PID)
```

LoadIVF *Load data file for IVFPred module*

Description

Load data file for IVFPred module

Usage

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

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by IVFPred
`UseExample`	chr. Method of uploading data. If "default",user should upload their own data files, or use "example#1" or "example#2" provided by this module.
`DataPath`	chr. Input directory of data file, e.g. "D:/test/eg_data_IVFPred.xlsx". It should be noted that the slash symbol is "/", not "\".
`VocaPath`	chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_IVFPred.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitIVF()
   res = LoadIVF(PID = res$PID, UseExample = "example#1")
```

`TransClassIVF` *Classify variables into various groups for IVF-ET data*

Description

Classify variables into various groups

Usage

```
TransClassIVF(PID, OutPath = "default", Vars, LevelTo)
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by initial functions.
`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.
`Vars`	Variables to be imputed. Available options include: "all.x", all exposure 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".
`LevelTo`	The number of levels to convert variables to.

Value

An R6 class object containing the variable(s) after classifying data into various levels.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitIVF()
    res1 = LoadIVF(PID = res$PID, UseExample = "example#1")
    res2 = TransClassIVF(PID = res$PID, Vars = "X47", LevelTo = 4)
    FuncExit(PID = res$PID)
```

TransDistrIVF *Transform variable distribution for IVF-ET data*

Description

Transform variable distribution

Usage

```
TransDistrIVF(PID,OutPath = "default", Vars, Method)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
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.
Vars	Variables to be imputed. Available options include: "all.x", all exposure 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".
Method	chr. Methods used for imputation. Available options include "lod" or "cart". For "lod" method, limit of detection (LOD) should be included in the "Vocabulary" file.

Value

An R6 class object containing the variable(s) after transforming distribution.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitIVF()
    res1 = LoadIVF(PID = res$PID, UseExample = "example#1")
    res2 = TransDistrIVF(PID = res$PID, Vars = "X61", Method = "log10")
    FuncExit(PID = res$PID)
```

`TransImputIVF` *Missing data imputation for IVF-ET data*

Description

Missing data imputation.

Usage

```
TransImputIVF(PID,OutPath = "default",Vars)
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by initial functions.
`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.
`Vars`	Variables to be imputed. Available options include: "all.x", all exposure 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".

Value

An R6 class object containing variable(s) with imputation.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitIVF()
    res1 = LoadIVF(PID = res$PID, UseExample = "example#1")
    res2 = TransImputIVF(PID=res$PID, Vars="all.x")
    FuncExit(PID = res$PID)
```

`TransScaleIVF` *Scale variables for IVF-ET data*

Description

Scale variables

Usage

```
TransScaleIVF(PID,OutPath = "default", Vars, Method = "normal",Direct="positive",
      RangeLow="0", RangeUpper="1")
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by initial functions.
`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.
`Vars`	Variables to be imputed. Available options include: "all.x", all exposure 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".
`Method`	chr. Scaling methods. Available options include "normal" and "range".
`Direct`	chr. Direction to be transformed, Available options include "positive" and "negative".
`RangeLow`	num. Lower limit for range method.
`RangeUpper`	num. Upper limit for range method. It should be greater than the lower limit.

Value

An R6 class object containing the variable(s) after scaling data.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitIVF()
    res1 = LoadIVF(PID = res$PID, UseExample = "example#1")
    res2 = TransScaleIVF(PID = res$PID, Vars = "X55", Method = "normal")
    FuncExit(PID = res$PID)
```

`TransTypeIVF` *Transform data type for IVF-ET data*

Description

Transform data type

Usage

```
TransTypeIVF(PID,OutPath = "default",Vars,To)
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by initial functions.
`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.
`Vars`	Variables to be imputed. Available options include: "all.x", all exposure 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".
`To`	chr. Indicate the type of the chosen variables to be transformed into. Available options include "integer", "numeric", "character", "factor", "logical", and "date".

Value

An R6 class object containing the variable(s) after transforming data type.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitIVF()
    res1 = LoadIVF(PID = res$PID, UseExample = "example#1")
    res2 = TransTypeIVF(PID = res$PID, Vars = "all.x", To = "character")
    FuncExit(PID = res$PID)
```

VizIVFPred *Visualize IVF-ET model results*

Description

VizIVFPred function is mainly aimed to visualize the modeling results calculated by IVFPred function. It can provide plots with high quality of the final results to make it easier for users to understand.

Usage

```
VizIVFPred(PID, OutPath = "default", Brightness = NULL, Palette = NULL)
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by initial functions.
`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.
`Brightness`	chr. Visualization brightness. Available options include "light" and "dark".
`Palette`	chr. Visualization palette. Available options include options about some journal preference styles including "cell", "nature", "science", "lancet", "nejm", etc.

Value

An R6 class object containing the model results

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
  res <- InitIVF()
 res1 = LoadIVF(PID = res$PID, UseExample = "example#2")
 res2 = TransImputIVF(PID=res$PID, Vars="all.x")
 res2 = DelLowVarIVF(PID=res$PID)
 res2 = DelMissIVF(PID=res$PID)
res2 = TransTypeIVF(PID = res$PID, Vars = "Y,X85,X86,X87,X88,X89,X90,X91,X92,X93,X94,X95,X96,X97,X98",
                     To = "factor")
 res2 = TransClassIVF(PID = res$PID, Vars = "X47", LevelTo = 4)
 res2 = TransScaleIVF(PID = res$PID, Vars = "X55", Method = "normal")
 res2 = TransDistrIVF(PID = res$PID, Vars = "X61", Method = "log10")
 res3 = IVFPred(PID = res$PID, AutTuneM = "random_search", AutTuneN = 5,
                RsmpMethod = "cv", Folds = 5,
                Learners = "logistic,rf,xgboost")
 res3 = VizIVFPred(PID = res$PID, Brightness = "light", Palette = "lancet")
 FuncExit(PID = res$PID)
```

VizIVFValid *Visualize IVF-ET validation results*

Description

VizIVFValid function is mainly aimed to visualize the modeling results calculated by IVFValid function. It can provide plots with high quality of the final results to make it easier for users to understand.

Usage

```
VizIVFValid(PID, OutPath = "default", Brightness = NULL, Palette = NULL)
```

Arguments

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

`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.

`Brightness` chr. Visualization brightness. Available options include "light" and "dark".

`Palette` chr. Visualization palette. Available options include options about some journal preference styles including "cell", "nature", "science", "lancet", "nejm", etc.

Value

An R6 class object containing the model results

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
  res <- InitIVF()
 res1 = LoadIVF(PID = res$PID, UseExample = "example#2")
 res2 = TransImputIVF(PID=res$PID, Vars="all.x")
 res2 = DelLowVarIVF(PID=res$PID)
 res2 = DelMissIVF(PID=res$PID)
res2 = TransTypeIVF(PID = res$PID, Vars = "Y,X85,X86,X87,X88,X89,X90,X91,X92,X93,X94,X95,X96,X97,X98",
                     To = "factor")
 res2 = TransClassIVF(PID = res$PID, Vars = "X47", LevelTo = 4)
 res2 = TransScaleIVF(PID = res$PID, Vars = "X55", Method = "normal")
 res2 = TransDistrIVF(PID = res$PID, Vars = "X61", Method = "log10")
 res3 = IVFPred(PID = res$PID, AutTuneM = "random_search", AutTuneN = 5,
                RsmpMethod = "cv", Folds = 5,
                Learners = "logistic,rf,xgboost")
 res4 = IVFValid(PID = res$PID, SingleGroup = TRUE)
 res4 = VizIVFValid(PID = res$PID, Brightness = "dark", Palette = "lancet")
 FuncExit(PID = res$PID)
```

Index