Package 'MSML'

November 21, 2023

Title Model selection based on Machine Learning (ML)

Version 1.0.0.0

Description Model selection based on Machine Learning (ML) approach following modified Recursive feature elimination (RFE) process.

License `use_mit_license()`, `use_gpl3_license()` or friends to pick a license

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.2

Depends R (>= 2.10)

LazyData true

R topics documented:

data_test 1
data_train 2
data_valid 3
model_combination 3
model_evaluation 4

Index **5**

`data_test` *7 sets of PRSs for test dataset and target phenotype*

Description

A dataset containing 7 sets of PRSs for test dataset and target phenotype

Usage

```
data_test
```

Format

A data frame for test dataset:

V1 PRS1, for bin1

V2 PRS2, for bin1

V3 PRS3, for bin1

V4 PRS4, for bin1

V5 PRS5, for bin1

V6 PRS6, for bin1

V7 PRS7, for bin1

target Target Phenotype, value

---

data_train *7 sets of PRSs for training data set and target phenotype*

Description

A dataset containing 7 sets of PRSs for training data set and target phenotype

Usage

```
data_train
```

Format

A data frame for training dataset:

V1 PRS1, for bin1

V2 PRS2, for bin1

V3 PRS3, for bin1

V4 PRS4, for bin1

V5 PRS5, for bin1

V6 PRS6, for bin1

V7 PRS7, for bin1

target Target Phenotype, value

`data_valid` *7 sets of PRSs for validation dataset and target phenotype*

Description

A dataset containing 7 sets of PRSs for validation dataset and target phenotype

Usage

```
data_valid
```

Format

A data frame for validation dataset:

V1 PRS1, for bin1

V2 PRS2, for bin1

V3 PRS3, for bin1

V4 PRS4, for bin1

V5 PRS5, for bin1

V6 PRS6, for bin1

V7 PRS7, for bin1

target Target Phenotype, value

`model_combination` *model_combination function This function will generate PRS based on all possible combianations of model. The total number of models required to explore the combinations of these 'n' features can be calculated by summing the combinations for each possible number of features, ranging from 1 to 'n' (?C(n,i)). where C(n,k) represents the binomial coefficient or "n choose k," with n denoting the total number of features and k indicating the number of features to include in each model.*

Description

model_combination function This function will generate PRS based on all possible combianations of model. The total number of models required to explore the combinations of these 'n' features can be calculated by summing the combinations for each possible number of features, ranging from 1 to 'n' (?C(n,i)). where C(n,k) represents the binomial coefficient or "n choose k," with n denoting the total number of features and k indicating the number of features to include in each model.

Usage

```
model_combination(data_train, data_valid, data_test, mv)
```

Arguments

data_train	This is the matrix for training dataset
data_valid	This is the matrix for validation dataset
data_test	This is the matrix for test dataset
mv	The total number of columns in data_train/data_valid/data_test

Value

This function will generate all possible model outcomes for validation and test dataset

Examples

```
data_train <- data_train
data_valid  <- data_valid
data_test  <- data_test
mv=8
model_combination(data_train,data_valid,data_test,mv)
```

---

model_evaluation *model_evaluation function*

Description

This function will identify best model in validation and test dataset.

Usage

```
model_evaluation(dat, mv, tn, prev)
```

Arguments

dat	This is the matrix for all the combinations of model
mv	The total number of columns in data_train/data_valid/data_test
tn	The total no of best models to be identified
prev	The prevalance of disease in the data

Value

This function will generate all possible model outcomes for validation and test dataset

Examples

```
dat <- read.table("models_test_all")
mv=8
tn=15
prev=0.047
model_evaluation(dat,mv,tn,prev)
```

Index

∗Topic **All**
model_combination, 3
∗Topic **Identify**
model_evaluation, 4
∗Topic **best**
model_evaluation, 4
∗Topic **combinations**
model_combination, 3
∗Topic **datasets**
data_test, 1
data_train, 2
data_valid, 3
∗Topic **models**
model_evaluation, 4
∗Topic **model**
model_combination, 3
∗Topic **possible**
model_combination, 3

C(n,i), *3*

data_test, 1
data_train, 2
data_valid, 3

model_combination, 3
model_evaluation, 4